# Huiswerk Wiskunde voor Neurale Netwerken
## Inleveren 24 september om 11:00

1. (a) Vele varianten mogelijk, zolang maar geldt:

$$\begin{aligned} w_1 * 0 + w_2 * 0 <& \ \theta \\ w_1 * 1 + w_2 * 0 \geq& \ \theta \\ w_1 * 0 + w_2 * 1 \geq& \ \theta \\ w_1 * 1 + w_2 * 1 \geq& \ \theta \end{aligned}$$

Oftewel:

$$\begin{aligned} &\theta > 0 \\ &w_1 \geq \theta \\ &w_2 \geq \theta \end{aligned}$$

De vierde eis volgt automatisch uit de eis dat $w_1$ en $w_2$ groter zijn dan $\theta$. Een voorbeeld van een goed antwoord is:

$$\theta = 0.5, w_1 = 1, w_2 = 1$$

(b) Er moet gelden:

$$\begin{aligned} &\theta > 0 \\ &w_1 < \theta \\ &w_2 < \theta \\ &w_1 + w_2 \geq \theta \end{aligned}$$

Voor $\theta = 0.5$ hebben we bijvoorbeeld: $w_1 = 0.4$ en $w_2 = 0.3$. Voor $\theta = -0.5$ zijn er geen oplossingen. (Immers, voor input (0,0) zou altijd 1 als output worden gegeven, wat niet mag).

(c) Ieder netwerk met $n$ inputs (zoals gegeven in de vraag) met willekeurige positieve gewichten is goed zolang geldt:

$$\theta \leq \sum_{i=1}^{n} w_i < (\theta + min(\vec{w}))$$

Waar $min(\vec{x})$ per definitie het kleinste element uit vector $\vec{x}$ is. Een voorbeeld is:

$$w_1 = w_2 = \cdots = w_n = 1 \text{ en } \theta = n - 0.5.$$

2. (a)

$$\begin{aligned}\sqrt{((\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2)} &= \\ \sqrt{(\tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4})} &= \\ \sqrt{1} &= \\ 1\end{aligned}$$

(b)

$$\frac{1}{2} * 1 + \frac{1}{2} * 0 + \frac{1}{2} * 1 + \frac{1}{2} * 0 = 1$$

(c)

$$\vec{v_3} = \vec{v_2} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

(d) Antwoord op de vraag: deze vector bestaat niet, er zijn vectoren $\vec{v_4}$ die willekeurige negatieve inproducte met $\vec{v_2}$ hebben. Stel

$$\vec{v_4} = \begin{pmatrix} -n \\ 0 \\ -n \\ 0 \end{pmatrix}$$

Antwoord op de bedoelde vraag (met $||\vec{v_4}|| = ||\vec{v_2}||$):

$$\vec{v_4} = -\vec{v_2} = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

(e) Inproduct het grootst: Als de vectoren naar dezelfde richting wijzen.

Inproduct het kleinst: Als de vectoren in tegenovergestelde richting wijzen.

Inproduct 0 (niet gevraagd overigens): Als de vectoren loodrecht op elkaar staan. Bijvoorbeeld (1,1) en (-1,1) of (1,-1,2) en (-2,0,1).

3. Gegeven zijn de matrix $A$ en de vectoren $x$ en $y$:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 4 & 0 & 0 \end{pmatrix}, \vec{x} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \vec{y} = \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$$

(a)

$$A^T = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 0 \\ 1 & 3 & 0 \end{pmatrix}$$

(b)

$$A\vec{y} = \begin{pmatrix} 1 * -1 + 2 * 0 + 1 * 4 \\ 3 * -1 + 2 * 0 + 3 * 4 \\ 4 * -1 + 0 * 0 + 0 * 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ -4 \end{pmatrix}$$

(c)

$$\begin{array}{cc} <\vec{x}, A\vec{y}> & = \\ \vec{x}^T(A\vec{y}) & = \\ \begin{pmatrix} 1 & 2 & -1 \end{pmatrix} A\vec{y} & = \\ \begin{pmatrix} 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 9 \\ -4 \end{pmatrix} & = \\ 25 & \end{array}$$

(d) Methode 1:

$$\begin{array}{cc} <\vec{x}, A\vec{y}> & = \\ \vec{x}^T(A\vec{y}) & = \\ (\vec{x}^T A)\vec{y} & = \\ <(\vec{x}^T A)^T, \vec{y}> & = \\ <A^T\vec{x}, \vec{y}> & \end{array}$$

Of, methode 2:

$$\begin{aligned}
A^T\vec{x} &= \\
\begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 0 \\ 1 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} &= \\
\begin{pmatrix} 3 \\ 6 \\ 7 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
<A^T\vec{x}, \vec{y}> &= \\
\begin{pmatrix} 3 & 6 & 7 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} &= \\
25 &= \\
<\vec{x}, A\vec{y}>
\end{aligned}$$

Het mag duidelijk zijn dat methode 1 (de algebraische methode) de voorkeur heeft. Niet alleen is het veel minder schrijfwerk, ook is het een bewijs dat geldt voor alle invullingen van $A, \vec{x}, \vec{y}$. De tweede methode laat zien dat het geldt voor deze specifieke vectoren en matrix, maar hieruit mag zeker niet geconcludeerd worden dat

$$<\vec{x}, A\vec{y}> = <A^T\vec{x}, \vec{y}>$$

algemeen geldig is, wat dus wel het geval is.

(e) Laat $r_i$ staat voor rij $i$ en $r_i \leftarrow x$ voor: "vervang rij $i$ door $x$".

$$(A \ : \ I)$$

$$\begin{pmatrix} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 3 & 2 & 3 & : & 0 & 1 & 0 \\ 4 & 0 & 0 & : & 0 & 0 & 1 \end{pmatrix}$$

$$r_2 \leftarrow r_2 - 3r_1$$

$$\begin{pmatrix} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & -4 & 0 & : & -3 & 1 & 0 \\ 4 & 0 & 0 & : & 0 & 0 & 1 \end{pmatrix}$$

$$r_3 \leftarrow r_3 - 4r_1$$

$$\begin{pmatrix} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & -4 & 0 & : & -3 & 1 & 0 \\ 0 & -8 & -4 & : & -4 & 0 & 1 \end{pmatrix}$$

$$r_3 \leftarrow r_3 - 2r_2$$

$$\begin{pmatrix} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & -4 & 0 & : & -3 & 1 & 0 \\ 0 & 0 & -4 & : & 2 & -2 & 1 \end{pmatrix}$$

$$r_3 \leftarrow -\frac{1}{4} r_3$$

$$r_2 \leftarrow -\frac{1}{4} r_2$$

$$\begin{pmatrix} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & 3/4 & -1/4 & 0 \\ 0 & 0 & 1 & : & -1/2 & 1/2 & -1/4 \end{pmatrix}$$

$$r_1 \leftarrow r_1 - 2r_2$$

$$\begin{pmatrix} 1 & 0 & 1 & : & -1/2 & 1/2 & 0 \\ 0 & 1 & 0 & : & 3/4 & -1/4 & 0 \\ 0 & 0 & 1 & : & -1/2 & 1/2 & -1/4 \end{pmatrix}$$

$$r_1 \leftarrow r_1 - r_3$$

$$\begin{pmatrix} 1 & 0 & 0 & : & 0 & 0 & 1/4 \\ 0 & 1 & 0 & : & 3/4 & -1/4 & 0 \\ 0 & 0 & 1 & : & -1/2 & 1/2 & -1/4 \end{pmatrix}$$

Dus $A^{-1}$:

$$\begin{pmatrix} 0 & 0 & 1/4 \\ 3/4 & -1/4 & 0 \\ -1/2 & 1/2 & -1/4 \end{pmatrix}$$

(f)

$$AA^{-1} = I$$

(Berekening hier weggelaten, werd wel verwacht bij het antwoord).

$$(AA^{-1})^{-1} = I^{-1} = I$$

$$(AA^{-1})^T = I^T = I$$

4. (a) Volgt uit definitie. Bedoeling was te laten zien dat je snapt wat er wordt bedoeld.

   (b) Het handigst is om eerst de XOR uit te schrijven:

| $x_1$ | $x_2$ | $z$ |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 1 | 0 |

   Als je $(x_1 \wedge x_2)$ defineert als 1 als $x_1$ en $x_2$ beiden waar zijn en anders 0 (ipv met true en false) en idem voor $\vee$, dan geldt ($\oplus$ staat voor de XOR operator):

$$x_1 \oplus x_2 = (x_1 \vee x_2) - (x_1 \wedge x_2)$$

   Laat een van de tussenliggende neuronen de waarde $y_1 = x_1 \vee x_2$ and de andere de waarde $y_2 = x_1 \wedge x_2$ berekenen. Dan kan het output neuron de waarde $y_1 - y_2$ (als een lineaire functie met gewichten $w_1 = 1$ en $w_2 = -1$) berekenen.

   Antwoord: Laat $\theta^h_a$ de threshold van het tussenliggende neuron nummer $a$ zijn en $\theta^o$ de threshold van de output. Verder is $w^{i\to h}_{ab}$ het gewicht tussen input $a$ en hidden neuron $b$ en $w^{h\to o}_a$ het gewicht tussen het hidden neuron $a$ en de output. Dan is een volgende invulling een goede oplossing:

$$\begin{aligned}
&w^{i\to h}_{11} = w^{i\to h}_{21} = 1\\
&w^{i\to h}_{12} = w^{i\to h}_{22} = 1\\
&w^{h\to o}_{1} = 1\\
&w^{h\to o}_{2} = -1\\
&\theta^h_1 = 0.5\\
&\theta^h_2 = 1.5\\
&\theta^o = 0.5
\end{aligned}$$

   (Meerdere oplossingen natuurlijk mogelijk).

   (c) Uit de bovenstaande oplossing blijkt eigenlijk al dat de threshold voor de output niet nodig is. De output van het netwerk zal hetzelfde zijn als de threshold gewoon niet wordt toegepast op die output. Het is niet mogelijk alle thresholds weg te halen (een bewijs hiervoor laat ik hier achterwegen).