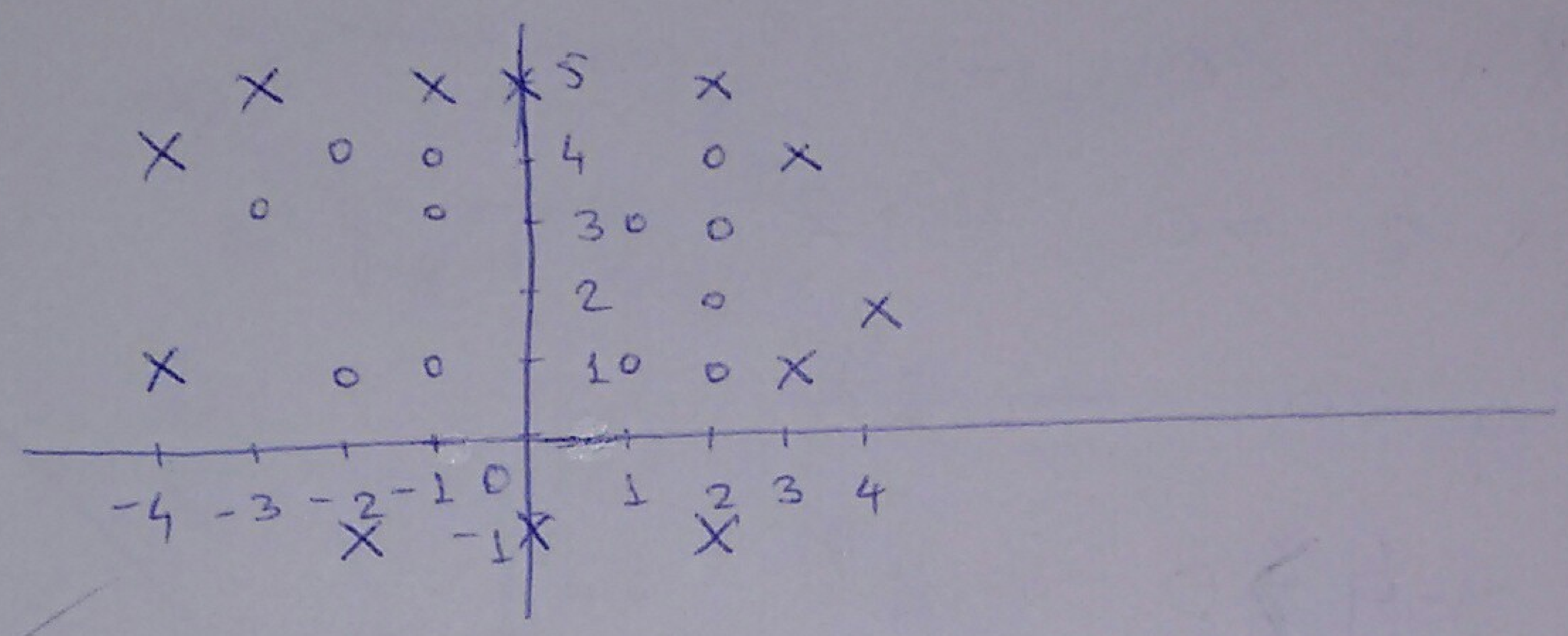
Κανονική 2016

Θέμα 1ο

α) 2-Δ χώρος → ο: κατηγορία 1, x: κατηγορία 2.



~~Για την πρώτη κατηγορία:~~ ~~$x_1 \in [-3, 2]$~~

~~$y_1 \in [0, 4]$~~

~~Για τη $2^η$ κατηγορία:~~ ~~$x_2 \in [-4$ … $[3,4]$~~

Παρατηρούμε ότι οι δύο κατηγορίες μπορούν να διαχωριστούν πλήρως με ένα ορθογώνιο, συγκρίνοντας τις x, y συντεταγμένες τους.

Συγκεκριμένα:

$(x<3$ ~~ή~~ $x>-4)$ κ' $(y>0$ ~~ή~~ $y<5) \to 1^η$

$x<-3$ ή $x>2$ ή $y<0$ ή $y>4 \to 2^η$

Θα σχεδιάσουμε ένα νευρωνικό δίκτυο δύο εξόδων, με έξοδο 10 αν το δεδομένο ανήκει στην $1^η$ κατηγορία κ' 01 αν το δεδομένο ανήκει στη $2^η$ κατηγορία.

~~Θα χρησιμοποιήσουμε τη δίτιμη απαίτηση.~~

Για κάθε μη έγκυρη είσοδο → έξοδος: 00

Επίσης, το ΝΔ θα είναι 2 εισόδων (x, y).

Πρέπει να σχεδιάσουμε τους εξής κανόνες:

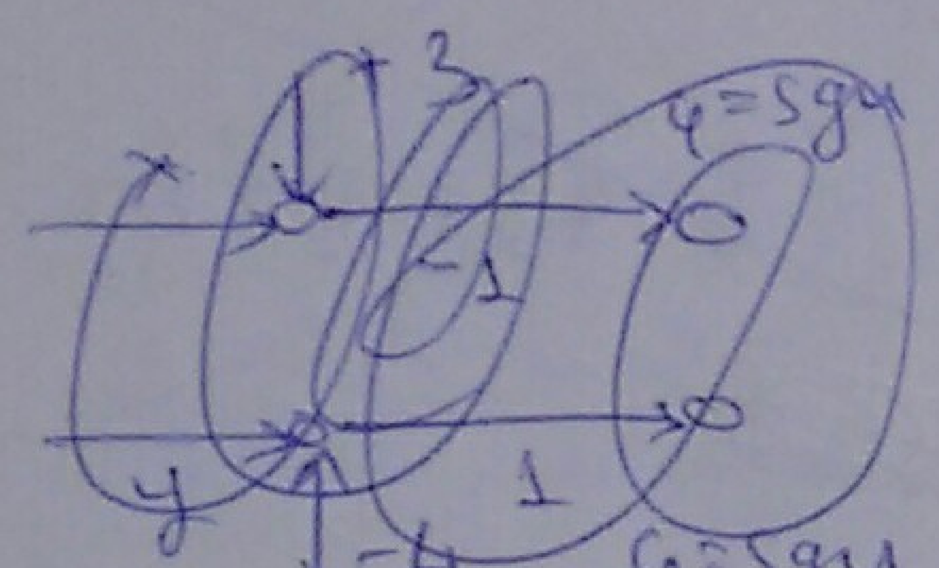
$x+3<0 \rightarrow 2^{\eta}$

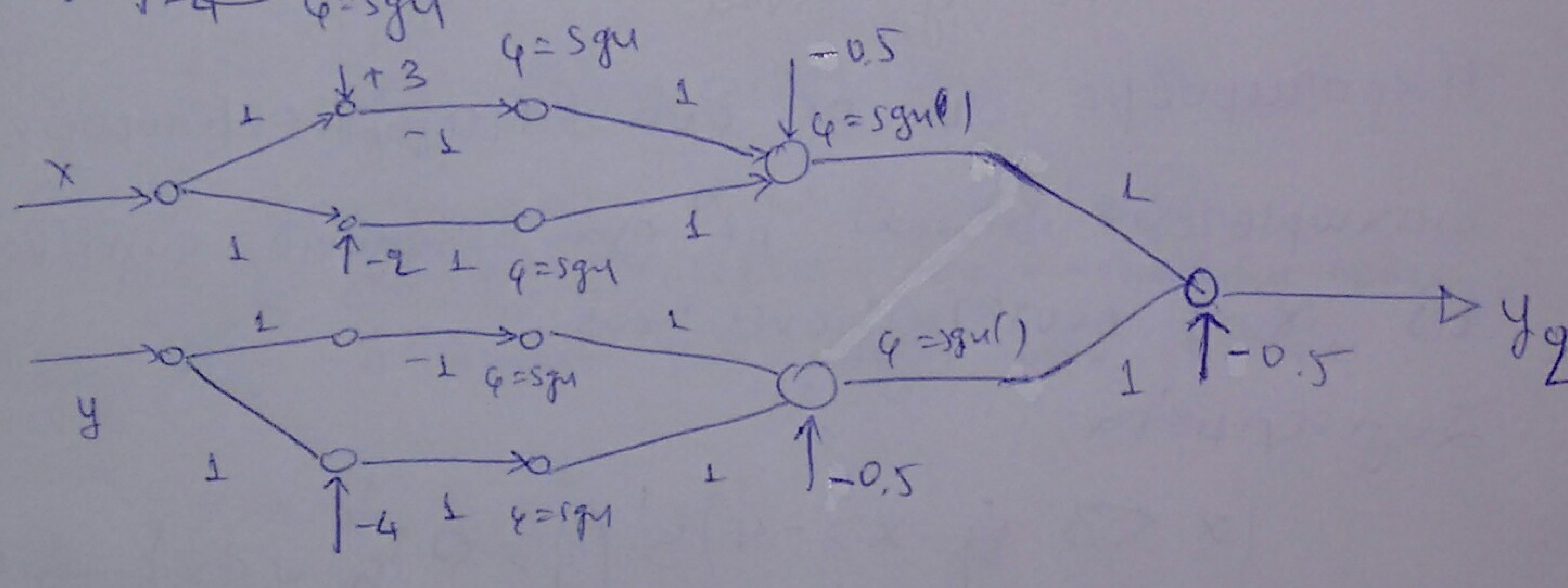
$x-2>0 \rightarrow 2^{\eta}$

$y<0 \rightarrow 2^{\eta}$

$y-4>0 \rightarrow 2^{\eta}$

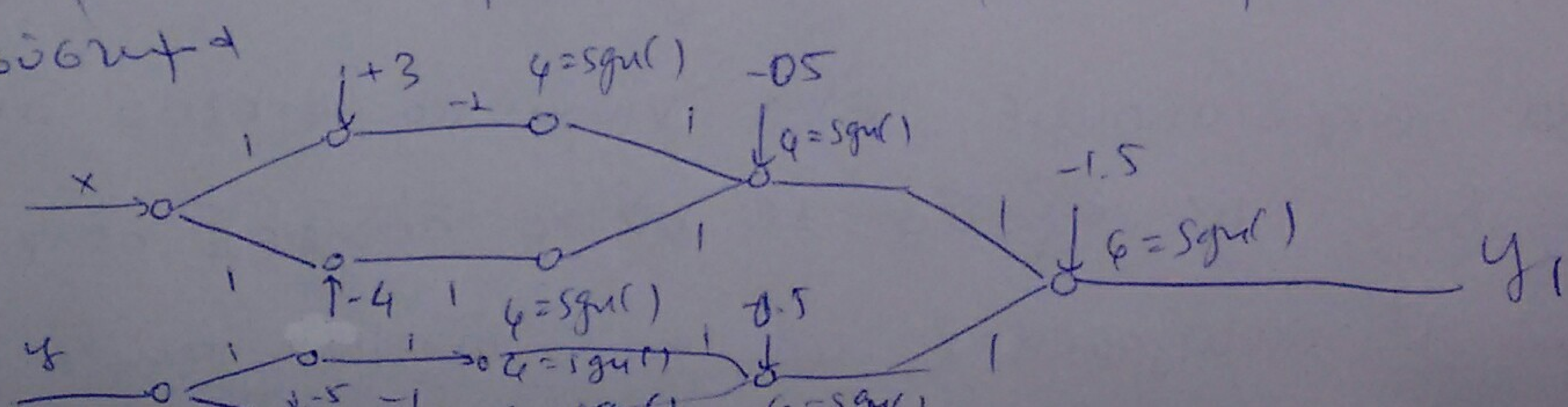
Υποσύστημα:



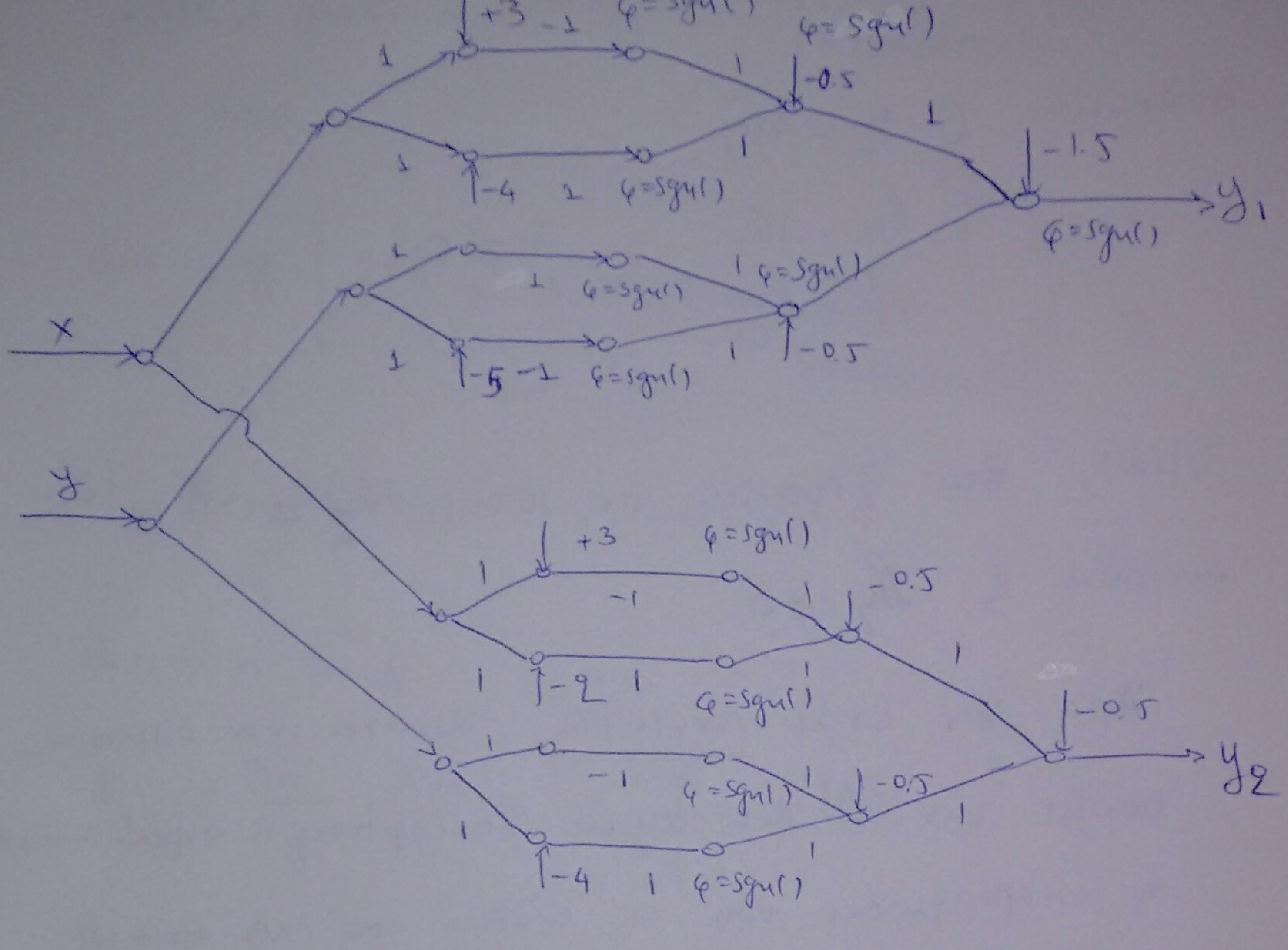


$(x-3<0$ ή $x+4>0)$ ή $(y>0$ ή $y-5<0)$

Υποσύστημα



Συνολικά



(β) Το 1^{o} perceptron έχει ορίσει μία ευθεία της μορφής $\alpha_1 x + \beta_1 y + \gamma_1 = 0$
(μιλάμε για απλό perceptron, άρα διαχωρίσιμα πρότυπα)
Το 2^{o} θέλουμε να ορίζει την ευθεία:

$$\alpha_2 x + \beta_2 y + \gamma_2 = 0$$

$$\rightarrow y = -\frac{\alpha_2}{\beta_2} x - \frac{\gamma_2}{\beta_2} \quad , \quad \beta_2 \neq 0$$

$$\& \quad -\frac{\gamma_2}{\beta_2} = 1$$

$$\rightarrow \beta_2 = -\gamma_2$$

Συνεπώς, 1η απαίτηση είναι να ισχύει $b_2 \neq 0$

Επιπλέον, για να ορίζεται η ίδια ευθεία,

θέλουμε $\begin{cases} a_2 = \lambda a_1 \\ b_2 = \lambda b_1 \\ \gamma_2 = \lambda \gamma_1 \end{cases}$, $\lambda \neq 0$

Όμως, δεν ξέρουμε αν ισχύει $b_1 \neq 0$, ϗ συνεπώς ϗ $b_2 \neq 0$

Άρα, δεν είναι πάντοτε δυνατό να υλοποιήσουμε την f με άλλο perceptron, αφού υπάρχει η περίπτωση ~~τα~~ η ευθεία που έχει οριστεί να είναι κατακόρυφη

$$x \to \circ \xrightarrow{w} \circ \xrightarrow{\varphi} y$$

γ) (i) $E = \frac{1}{2}(d-y)^2 = \frac{1}{2}e^2$ $\qquad \varphi = \frac{1}{1+\exp(-av)}$

$\Delta w = \eta \cdot \delta_i \cdot y_i$

$$\delta_i(n) = -\frac{\partial E(n)}{\partial y_j(n)} \cdot \frac{\partial y_i(n)}{\partial v_i(n)} = -\frac{\partial E(n)}{\partial y(n)} \varphi_i'(v(n)) =$$

$$= -\frac{\partial E}{\partial e} \cdot \frac{\partial e}{\partial y} \cdot \varphi'(v) = -e \cdot 2(d-y)(-1) \cdot \left(-\frac{1}{(1+\exp(-av))^2}\right) \cdot (-a)\exp(-av) = 2(d-y)^2 a \exp(-av) \cdot \frac{1}{(1+\exp}$$

Συνεπώς, $\delta = 2(d-y)^2 a \exp(-av) \cdot \dfrac{1}{(1+\exp(-av))^2}$

ξ' άρα

$$\Delta w = \eta \cdot \left[2(d-y)^2 a \cdot \exp(-av) \cdot \frac{1}{(1+\exp(-av))^2} \right] \cdot y.$$

(ii) $E = -(d \log y + (1-d)\log(1-y))$

$$\Delta w = \eta \cdot \delta \cdot y.$$

$$\delta = -\frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial v} = -\frac{\partial E}{\partial y} \cdot \varphi'(v).$$

$$\delta \cancel{\frac{\partial E}{\partial y}} = -\left\{ \left[d \cdot \frac{1}{y} + (1-d) \cdot \frac{1}{1-y}(-1) \right] \right\} \cdot \varphi'(v)$$

$$= -\frac{1-d}{1-y} - \frac{d}{y} \cdot \varphi'(v) = \frac{y(1-d) - d(1-y)}{y(1-y)} \cdot \varphi'(v) =$$

$$= -\frac{y - yd - d + yd}{y(1-y)} \cdot \varphi'(v) = \frac{y-d}{y(1-y)} \cdot \varphi'(v) =$$

$$= -\frac{y-d}{y(1-y)} \cdot a\exp(-av) \cdot \frac{1}{(1+\exp(-av))^2}$$

Άρα

$$\Delta w = \eta \cdot \frac{d-y}{y(1-y)} \cdot a \exp(-av) \cdot \frac{1}{(1+\exp(-av))^2} \cdot \cancel{y}$$

Όμως, $y = \varphi(\upsilon) = \dfrac{1}{1+\exp(-\alpha\upsilon)} \Rightarrow$

$\cancel{\Rightarrow y + y\exp(-\alpha\upsilon) = 1 \Rightarrow}$

$\Rightarrow \exp(-\alpha\upsilon) = \dfrac{1}{y} - 1.$

Άρα για το (i): $\Delta w = \eta 2(d-y)^2 \cdot a \cdot \left(\dfrac{1}{y} - 1\right) \cdot y^2 \cdot y =$

$= \eta \cdot 2(d-y)^2 \cdot a \cdot (1-y) \cdot y^2$

Και για το (ii): $\Delta w = \eta \cdot \dfrac{d-y}{(1-y)\cdot \cancel{y}} \cdot a \cdot \left(\dfrac{1}{y} - 1\right) \cdot y^2 \cdot \cancel{y} =$

$= \eta \cdot \dfrac{d-y}{\cancel{1-y}} \cdot a \cdot \dfrac{(\cancel{1-y})}{\cancel{y}} \cdot y^{\cancel{2}} =$

$= \eta \cdot (d-y) \cdot a \cdot y$

Άρα, γραμμική συνάρτηση θα διευκόλυνε στο (ii), καθώς θα απλοποιούσε την εξίσωση ανανέωσης των βαρών.