MOE kapuzó hálózat tanítása

1. Egy 2 bemenetű 1 kimenetű (MISO) MOE struktúrában lineáris kapuzó hálózatot és két lineáris szakértőt használunk. A szakértők paraméterei rendre $\boldsymbol{\theta}_1=[1 \quad 2]^T$, $\boldsymbol{\theta}_2=[0 \quad 1]^T$, nincs eltolás (*bias*). A kapuzó hálózat megfelelő paramétervektorai $\mathbf{v}_1=[2 \quad 1]^T$, $\mathbf{v}_2=[1\ 4]^T$, itt sincs eltolás.
A szakértők kimenetéhez nulla várható értékű, Gauss zajt asszociálunk, az első szakértőnél $\sigma_1=2$, a másodiknál $\sigma_2=1$. A tanítás során a bátorsági faktor 0,1.

A. Adja meg a MOE struktúra kimeneti jelét, ha a bemeneti vektor $\mathbf{x}=[+3\ +1]^T$!

B. Adja meg a kapuzó hálózat $\mathbf{v}_2$ paramétervektorának új értékét! A tanítást az $\mathbf{x}^{(l)}=[+3\ +1]^T$ $y^{(l)}=0$ mintával végezzük!

C. Számítsa ki és adja meg, hogy a bemeneti paraméterek – x_1 és x_2 – által kifeszített sík mely tartományában lesz az első, illetve a második szakértőnek nagyobb súlya az eredő kimenet kialakításában!

D. Számítsa ki és adja meg, hogy a bemeneti paraméterek – x_1 és x_2 – által kifeszített sík mely tartományában lesz az első, illetve a második szakértőnek nagyobb súlya az eredő kimenet kialakításában, ha a két nemlineáris szakértő:

$$g_i = \frac{e^{f_i(\mathbf{v}_i^T\mathbf{x})}}{e^{f_1(\mathbf{v}_1^T\mathbf{x})}+e^{f_2(\mathbf{v}_2^T\mathbf{x})}} \qquad \xi_1 = f_1(\mathbf{v}_1^T\mathbf{x}) = sigm(\mathbf{v}_1^T\mathbf{x}) \quad \text{és} \quad \xi_2 = f_2(\mathbf{v}_2^T\mathbf{x}) = sigm(\mathbf{v}_2^T\mathbf{x})$$

E. Milyen lesz a két szakértő "felségterületét" elválasztó határvonal, ha nemlineáris szakértőket használunk:

$$g_i = \frac{e^{f_i(\mathbf{v}_i^T\mathbf{x})}}{e^{f_1(\mathbf{v}_1^T\mathbf{x})}+e^{f_2(\mathbf{v}_2^T\mathbf{x})}} \qquad \xi_1 = f_1(\mathbf{v}_1^T\mathbf{x}) = sigm(\mathbf{v}_1^T\mathbf{x}) \qquad \xi_2 = f_2(\mathbf{v}_2^T\mathbf{x}) = \sin(sigm(\mathbf{v}_2^T\mathbf{x}))$$

MOE képletek:

$$\frac{\partial L}{\partial \mu_i} = \frac{g_i^{(\ell)}}{\sum_j g_j^{(\ell)} P(y^{(\ell)} \mid x^{(\ell)}, \theta_j)} \frac{\partial}{\partial \mu_i} P(y^{(\ell)} \mid x^{(\ell)}, \theta_i) \quad ; \quad h_i^{(\ell)} = \frac{g_i^{(\ell)} P(y^{(\ell)} \mid x^{(\ell)}, \boldsymbol{\theta}_i)}{\sum_j g_j^{(\ell)} P(y^{(\ell)} \mid x^{(\ell)}, \boldsymbol{\theta}_j)}$$

Gauss asszociált zaj és lineáris szakértő esetén: $$\frac{\partial L}{\partial \boldsymbol{\theta}_i} = h_i^{(\ell)} \frac{(y^{(\ell)} - \boldsymbol{\mu}_i^{(\ell)})\mathbf{x}^{(\ell)}}{\sigma_i^2}$$

Lineáris kapuzó hálózat tanítási összefüggése: $$\frac{\partial L}{\partial \mathbf{v}_i} = \left(h_i^{(\ell)} - g_i^{(\ell)}\right)\mathbf{x}^{(\ell)}$$

MOE kapuzó hálózat tanítása

Példa (Matlab stílusban) A. és B. alkérdések

```
%MOE kapuzo halozat tanitas (ket szakerto)
% #1 szakerto:      T1=[1 2]';
                    mu1=0;  %asszocialt zaj
                    o1=2;
% #2 szakerto:      T2=[0 1]';
                    mu2=0;  %asszocialt zaj
                    o2=1;
%kapuzo halozat parameterei (linearis kapuzo halozat - nincs eltolas)
v1=[2 1]';
v2=[1 4]';
```

pl $\mathbf{x}^{(l)}=[0\ 1]^T$-re $\mathbf{v}_2{}'*\mathbf{x}^{(l)}>\mathbf{v}_1{}'*\mathbf{x}^{(l)}$

```
%tanito minta
xl=[3 1]';
yl=0;
%a tanitas batorsagi faktora
r=0.1;
%------------------------------
```

Megoldás

pylxl1=gaussprob(T1'*xl,o1,yl); = **0.0088**
pylxl2=gaussprob(T2'*xl,o2,yl); = **0.2420**

x1=v1'*xl; = **7**
x2=v2'*xl; = **7**

g1=exp(x1)/(exp(x1)+exp(x2)); = **0,5**
g2=1-g1; = **0,5**
sumgp=g1*pylxl1+g2*pylxl2; = **0,1254**

h1=g1*pylxl1/sumgp; = **0,0350**
h2=g2*pylxl2/sumgp; = **0.9650** (mivel g1=g2, ezért h1+h2=1)

dLdx1=h1-g1; = **-0,4650**
dLdx2=h2-g2; = **+0.4650**

$\Delta\mathbf{v}_1$ =r*dLdx1*xl; = **[-0,1395, -0,0465]**^{T}
$\Delta\mathbf{v}_2$ =r*dLdx2*xl; = **[+0,1395, +0,0465]**^{T}

%-----------------------------

A tanítás után az új a priori valószínűségek:
$(\mathbf{v}_1+\Delta\mathbf{v}_1)'*\mathbf{x}^{(l)}$= **6,5350**
$(\mathbf{v}_2+\Delta\mathbf{v}_2)'*\mathbf{x}^{(l)}$= **7,4650**
g_1 = exp(x1)/(exp(x1)+exp(x2)) = **0,2829**
$g_2=1-g_1$ = **0,7171**

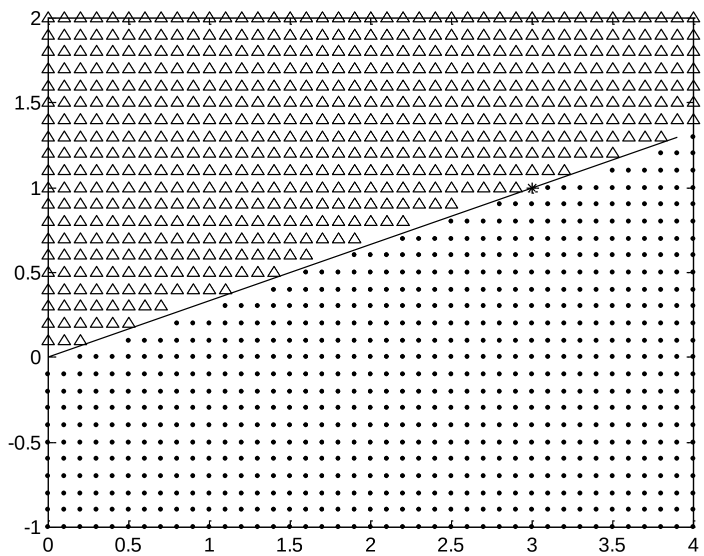
C. alkérdés

Mivel a két paramétervektor különbsége:

$\mathbf{v}_1=[2\ 1]^T$ $\mathbf{v}_2=[1\ 4]^T$ $\mathbf{v}_1-\mathbf{v}_2=[\,1, -3]^T$,

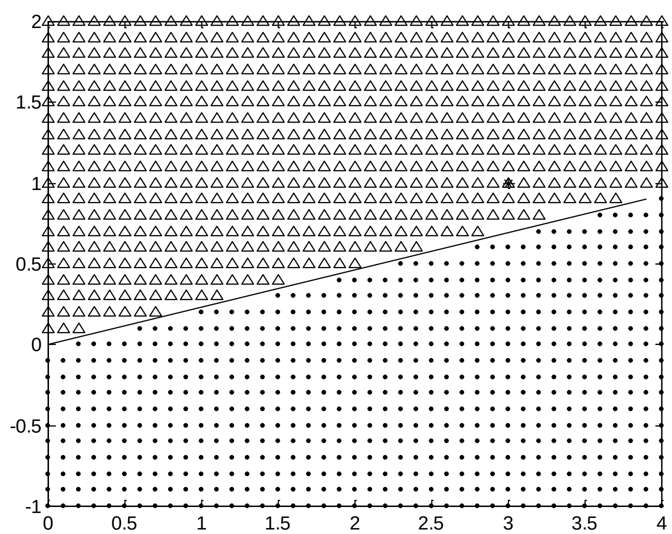
ezért a két szakértő által „uralt” terület elválasztó egyenese az y=x/3 egyenlettel adható meg.
Tanítás előtt: a felső részben a #2-es ($g_2>g_1$) az alsó részben az #1-es ($g_2<g_1$) szakértőben nagyobb a bizalom. ’*****’ jelöli a tanító pontot $\mathbf{x}^{(l)}=[3\ 1]^T$).

MOE kapuzó hálózat tanítása



Tanítás után: az új paraméter vektorok:

$\mathbf{v}_1+\mathbf{\Delta v}_1$= [1.8605, 0.9535]^T
$\mathbf{v}_2+\mathbf{\Delta v}_2$= [1.1395, 4.0465]^T
$(\mathbf{v}_1+\mathbf{\Delta v}_1)-(\mathbf{v}_2+\mathbf{\Delta v}_2)$= [0.7210, -3.0930]^T



D. alkérdés

A határvonal nem változik, mivel mindkét szakértőnél ugyanazt a nemlinearitást használjuk, végül ugyanarra az egyenletre jutunk, mint C. alatt, amikor mindkét szakértő lineáris volt.

E. alkérdés

a nemlineáris kapuzó hálózat paraméterei és struktúrája

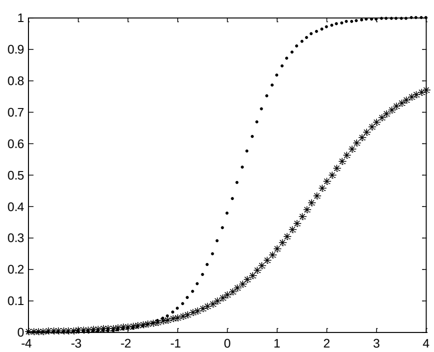
MOE kapuzó hálózat tanítása

$\mathbf{v}_1$=[2 1]'; $\mathbf{v}_2$=[1 4]';

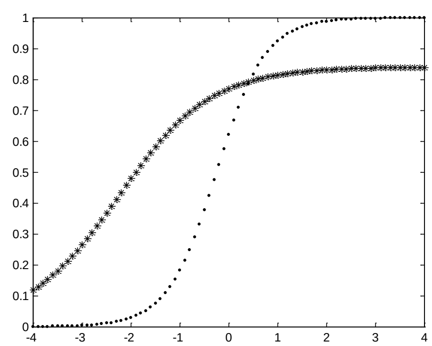
$$\xi_1 = \frac{1}{1+\exp(-\mathbf{v}_1\mathbf{x}^{(l)})} \qquad \xi_2 = \sin\left(\frac{1}{1+\exp(-\mathbf{v}_2\mathbf{x}^{(l)})}\right)$$

$$g_1 = \frac{\exp(x_1)}{\exp(\xi_1)+\exp(\xi_2)} \qquad g_2 = 1 - g_1$$

x2=-0,5 egyenletű vízszintes egyenes mentén ξ_1 és ξ_2



x2=+0,5 egyenletű vízszintes egyenes mentén ξ_1 és ξ_2



x2=+2 egyenletű vízszintes egyenes mentén ξ_1 és ξ_2