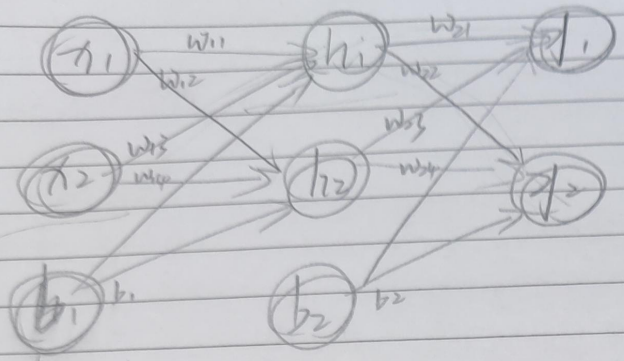
NO.
DATE



$w_{11}=0.24$ $w_{21}=0.5$
$w_{12}=0.3$ $w_{22}=0.45$
$w_{13}=0.2$ $w_{23}=0.55$
$w_{14}=0.16$ $w_{24}=0.4$
$b_1=0.28$ $b_2=0.6$
$x_1=0.08$ $x_2=0.12$
$label_1=0.05$ $label_2=0.95$
$\alpha=0.5$

forward:

$net_{h1}=w_{11}x_1+w_{13}x_2+b_1=0.24\times0.08+0.2\times0.12+0.28=0.3232$

$net_{h2}=w_{12}x_1+w_{14}x_2+b_1=0.3\times0.08+0.16\times0.12+0.28=0.3232$

$out_{h1}=1/(1+e^{-net_{h1}})=0.580104$

$out_{h2}=1/(1+e^{-net_{h2}})=0.580104$

$o_1=w_{21}\cdot out_{h1}+w_{23}\cdot out_{h2}+b_2$
$=1.209109$

$o_2=w_{22}\cdot out_{h1}+w_{24}\cdot out_{h2}+b_2$
$=1.093088$

$E_{total}=\sum\frac{1}{2}(target-output)^2=$ (E_1 E_2)
$=\frac{1}{2}[(1.209109-0.05)^2+(1.093088-0.95)^2]=0.671761+0.010223$
$=0.681984$

backward 输出层:

$\frac{\partial E_{total}}{\partial w_{21}}=\frac{\partial E_{total}}{\partial o_1}\cdot\frac{\partial o_1}{\partial w_{21}}=\frac{\partial\sum\frac{1}{2}(o_i-label_i)^2}{\partial o_1}\cdot\frac{\partial o_1}{\partial w_{21}}$

$=(o_1-label_1)\cdot out_{h1}=0.626404$

$w_{21}^+=w_{21}-\alpha\cdot\frac{\partial E_{total}}{\partial w_{21}}=0.1638$ 同理,可以求得

$w_{22}^+=0.4085$ $w_{23}^+=0.2138$ $w_{24}^+=0.3585$

对于 b_2 同样有 $\frac{\partial E_{total}}{\partial b_2}=\frac{\partial E_{total}}{\partial o_1}\cdot\frac{\partial o_1}{\partial b_2}+\frac{\partial E_{total}}{\partial o_2}\cdot\frac{\partial o_2}{\partial b_2}=(o_1+o_2-label_1-label_2)=1.302199$

$b_2^+=b_2-\alpha\cdot\frac{\partial E_{total}}{\partial b_2}=0.051$

NO.
DATE

backward 隐藏层.

$$\frac{\partial E_{total}}{\partial w_{11}} = \frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} \cdot \frac{\partial net_{h_1}}{\partial w_{11}}$$

而① $\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_1}{\partial out_{h_1}} + \frac{\partial E_2}{\partial out_{h_1}}$

又: $\frac{\partial E_1}{\partial out_{h_1}} = \frac{\partial E_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial out_{h_1}}$

$= (y_1 - label_1) \cdot w_{21}$

$\frac{\partial E_2}{\partial out_{h_1}} = \frac{\partial E_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial out_{h_1}}$

$= (y_2 - label_2) \cdot w_{22}$

② $\frac{\partial out_{h_1}}{\partial net_{h_1}} = \frac{\partial (1/(1+e^{-net_{h_1}}))}{\partial net_{h_1}} = out_{h_1}(1 - out_{h_1})$

③ $\frac{\partial net_{h_1}}{\partial w_{11}} = x_1$

故有 $\frac{\partial E_{total}}{\partial w_{11}} = [(y_1 - label_1) \cdot w_{21} + (y_2 - label_2) \cdot w_{22}] \cdot out_{h_1}(1 - out_{h_1}) \cdot x_1$

最终有 $w_{11}^+ = w_{11} - \partial \cdot \frac{\partial E_{total}}{\partial w_{11}} = 0.2338$ 同理.

$w_{12}^+ = 0.2438$ $w_{13}^+ = 0.1906$ $w_{14}^+ = 0.1899$

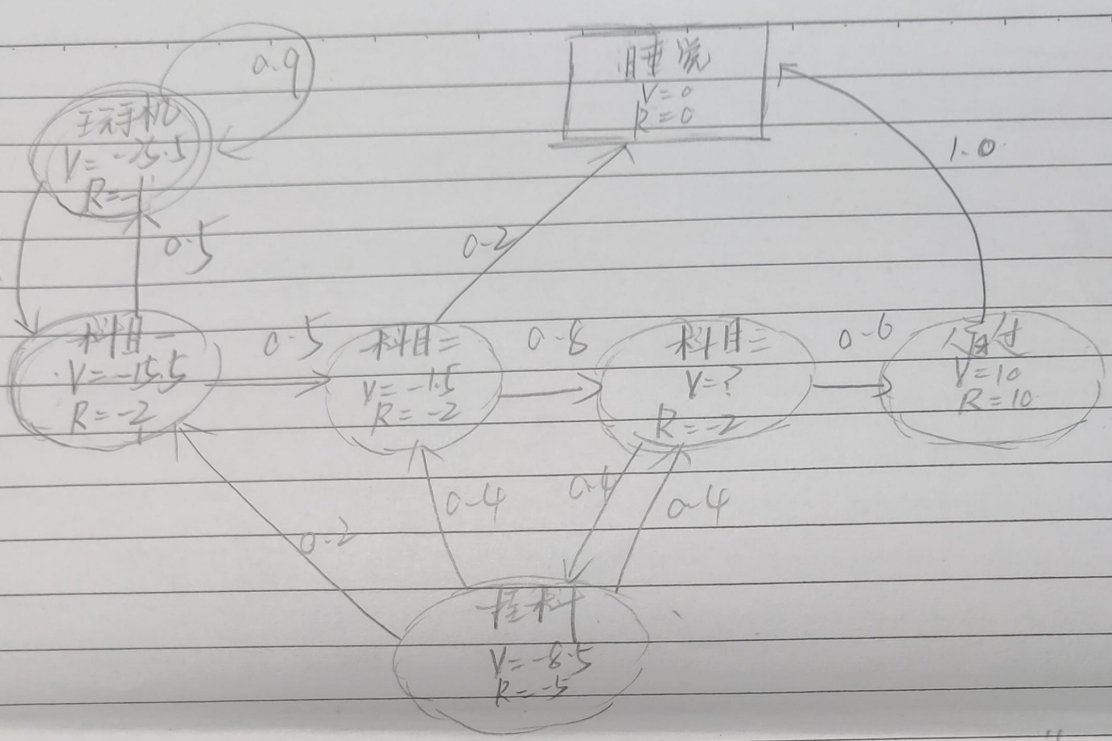
$b_1^+ = b_1 - \partial \cdot \frac{\partial E_{total}}{\partial b_1}$

而 $\frac{\partial E_{total}}{\partial b_1} = [(y_1 - label_1) \cdot w_{21} + (y_2 - label_2) \cdot w_{22}] \cdot out_{h_1}(1 - out_{h_1}) \cdot 1 +$

$[(y_1 - label_1) \cdot w_{23} + (y_2 - label_2) \cdot w_{24}] \cdot out_{h_2}(1 - out_{h_2}) \cdot 1$

得 $b_1^+ = 0.1117$

NO.
DATE



(1): $\gamma = 0.5$ ① 轨迹："科目一，科目二，科目三，通过，睡觉"

② 轨迹："科目一，玩手机，玩手机，科目一，科目二，睡觉"

① $V_t = -2 + (-2)\times 0.5 + (-2)\times 0.5^2 + 10\times 0.5^3 + 0\times 0.5^4$

$= -2.25$

② $V_t = -2 + (-1)\times 0.5 + (-1)\times 0.5^2 + (-2)\times 0.5^3 + (-2)\times 0.5^4 + 0\times 0.5^5$

$= -3.125$

(2): $\gamma = 1$, $V_{\pi}(s_t) = E_{A\sim\pi}[Q_{\pi}(s_t, A)] = E[U_t \mid s_t]$.

$= E[R_t + \gamma U_{t+1} \mid s_t]$

$= E[R_t \mid s_t = s_t] + \gamma E[U_{t+1} \mid s_t = s_t]$

$= R_s + \gamma \sum_{s'} P_{ss'} E[U_{t+1} \mid s_t = s_t, s_{t+1} = s']$

$= R_s + \gamma \sum_{s'} P_{ss'} E[U_{t+1} \mid s_{t+1} = s']$

由贝尔曼方程，$= R_s + \gamma \sum_{s'} P_{ss'} V(s')$，故算得 $V = -2 + 0.6\times 10 + 0.4\times -8.5$

$= 0.6$