2021

問1

(1-1) $g(x) = w^T x = 0.2 + 0.3x_1$

$x_1 = -0.2$ のデータは、$g(x)|_{x_1=-0.2} = 0.14 > 0$ となるため正しく識別できない。

(1-2)

以下 $g(x) = g([1, x_1]^T)$ のことを $g(x_1)$ とかく。

$g(1.0) = 0.5 > 0$ より ok

$g(0.5) = 0.35 > 0$ より ok

$g(-0.2) = 0.14 > 0$ より 更新

$w' = [0.2, 0.3] - 0.5[1, -0.2] = [-0.3, 0.4]$

$g(x) = -0.3 + 0.4x_1$

$g(-1.3) = -0.82 < 0$ より ok

$g(1.0) = 0.1 > 0$ より ok

$g(0.5) = -0.1 < 0$ より 更新

$w' = [-0.3, 0.4] + 0.5[1, 0.5] = [0.2, 0.65]$

$g(x) = 0.2 + 0.65x_1$

$g(-0.2) = 0.07 > 0$ より 更新

$w' = [0.2, 0.65] - 0.5[1, -0.2] = [-0.3, 0.75]$

$g(x) = -0.3 + 0.75x_1$

$g(-1.3) < 0$ より ok

$g(1.0) = 0.45 > 0$ より ok

$g(0.5) = 0.075 > 0$ より ok

$g(-0.2) < 0$ より ok

$\therefore g(x) = -0.3 + 0.75x_1$

(1-3)
識別関数の重みの変化量が大きくなり、重みの初期値が終了条件領域から遠く離れている場合には素早く、終了条件領域内に更新される可能性があるが、同時に、初期値によっては終了条件領域から離れてしまい、非常に時間がかかってしまう可能性がある。

(1-4)
入力ベクトルに対する誤差の二乗から得られる損失関数Jが最小化されたときに終了する。
更新はJの重み方向の勾配 εx を用い、
$w' = w - \rho \varepsilon x$ によって与えられる。

(1-5)
出力層にて与えられる教師信号は、中間層では直接利用できないから。
バックプロパゲーション
出力結果と教師信号との誤差を出力層側から入力層側へと伝播させていくことで実質的に中間層でも教師信号を利用できるようにし、重みの学習を可能とする方法。

(1-6)
交差検証は、教師あり学習でないと行えず、図1の例では教師なし学習であるから。

(2-1)

ドメイン：DB群の中からどのDBに対する問題なのかを決定する.

意図：文章内の各要素がDB内のどの属性に対応しているかを決定する

スロット：意図～で決定した属性に従ってDBから値を取り出す。

(2-2)

スロット値同定

ドメイン～，意図～は多クラス分類であるが，スロット～は系列ラベリング問題である。

(2-3)

信号の振幅の絶対値が一定以下であるような区間の長さが一定以上あるとき，ここを発話の区切りとする

意味的なまとまりを一発話とする

話者の交代したときを発話の区切りとする。

話者交代～では，話者が交代したときが発話の区切りであるが，システムは，ユーザの発話が終わってからでないと発話することができないため。

(3-1)

$$J(\boldsymbol{W}) = \frac{1}{T}\sum_{t=1}^{T} |\boldsymbol{W}^T \boldsymbol{s}_t - d_t|^2 = \frac{1}{T}\sum_{t=1}^{T} (\boldsymbol{W}^T \boldsymbol{s}_t - d_t)(\boldsymbol{W}^T \boldsymbol{s}_t - d_t)^T$$

(3-2)

$$\begin{aligned}
J(\boldsymbol{W}) &= E[|\boldsymbol{W}^T \boldsymbol{s}_t - d_t|^2] \\
&= E[(\boldsymbol{W}^T \boldsymbol{s}_t - d_t)(\boldsymbol{s}_t^T \boldsymbol{W} - d_t)] \\
&= E[\boldsymbol{W}^T \boldsymbol{s}_t \boldsymbol{s}_t^T \boldsymbol{W} - d_t \boldsymbol{W}^T \boldsymbol{s}_t - d_t \boldsymbol{s}_t^T \boldsymbol{W} + d_t^2] \\
&= E[d_t^2] - 2\boldsymbol{W}^T E[d_t \boldsymbol{s}_t] + \boldsymbol{W}^T E[\boldsymbol{s}_t \boldsymbol{s}_t^T] \boldsymbol{W}
\end{aligned}$$

$$\frac{\partial J}{\partial \boldsymbol{W}} = -2E[d_t \boldsymbol{s}_t] + 2E[\boldsymbol{s}_t \boldsymbol{s}_t^T]\boldsymbol{W}$$

$\frac{\partial J}{\partial \boldsymbol{W}} = 0$ のとき．誤差が最小になると考えれ．

$$E[\boldsymbol{s}_t \boldsymbol{s}_t^T]\boldsymbol{W} = E[d_t \boldsymbol{s}_t]$$

$$\therefore \boldsymbol{W} = (E[\boldsymbol{s}_t \boldsymbol{s}_t^T])^{-1} E[d_t \boldsymbol{s}_t]$$

(3-3)

$$\begin{aligned}
\boldsymbol{W}_k &= \boldsymbol{W}_{k-1} - \mu \nabla J(\boldsymbol{W}_{k-1}) \\
&= \boldsymbol{W}_{k-1} + 2\mu(E[d_t \boldsymbol{s}_t] - E[\boldsymbol{s}_t \boldsymbol{s}_t^T]\boldsymbol{W}_{k-1})
\end{aligned}$$

(3-4)

$$\boldsymbol{W}_{t+1} = \boldsymbol{W}_t - \alpha \boldsymbol{s}_t d_t$$