3) a) Vorlesung K4 // SL 17

$$l\big((w; x_n, t)_{n=1,\dots,5}\big)$$
$$= \sum_{n=1}^{5} \Big(t_n \, w^T \phi(x_n) - \ln\big(1+\exp(w^T\phi(x_n))\big) \Big)$$

✓

An den Trainingsdatensatz

$x_1 \; x_2 \; t$

$$\{(1,3,0),(2,4,0),(4,1,0),(3,1,1),(4,2,1) \in \mathbb{R}^2 \times \{0,1\}\}\,.$$

1 1 1

$$= \sum_{n=1}^{5} \Big(t_n(w_0 + w_1x_1 + w_2x_2) - \ln\big(1+\exp(w_0+w_1x_1+w_2x_2)\big)\Big)$$
$$= -\ln\big(1+\exp(w_0 + w_1 + 3w_2)\big) - \ln\big(1+\exp(w_0+2w_1+4w_2)\big)$$
$$- \ln\big(1+\exp(w_0+4w_1+w_2)\big) + (w_0+3w_1+w_2)$$
$$- \ln\big(1+\exp(w_0+3w_1+w_2)\big) + (w_0+4w_1+2w_2)$$
$$- \ln\big(1+\exp(w_0+4w_1+2w_2)\big)$$

b) Iterated Reweighted Least Squares (IRLS)

$$\text{II} \quad w^{(\tau+1)} = \big(\Phi^T R(w^{(\tau)})\Phi\big)^{-1} \cdot$$
$$\Phi R(w^{(\tau)}) z(w^{(\tau)})$$

$$z(w^{(\tau)}) = \Phi w^{(\tau)} - R(w^{(\tau)})]^{-1} \cdot (\pi(w^{(\tau)}) - t)$$

An den Trainingsdatensatz

$$\{\overset{x_1}{(1,3,0)}, \overset{x_1}{(2,4,0)}, \overset{x_1}{(4,1,0)}, \overset{x_1}{(3,1,1)}, \overset{x_1}{(4,2,1)} \in \mathbb{R}^2 \times \{0,1\}\}\,.$$

$$\mathbf{\Phi} = \begin{pmatrix} \phi_0(\mathbf{x}_1) & \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) \\ \phi_0(\mathbf{x}_2) & \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) \\ \phi_0(\mathbf{x}_3) & \phi_1(\mathbf{x}_3) & \phi_2(\mathbf{x}_3) \\ \phi_0(\mathbf{x}_4) & \phi_1(\mathbf{x}_4) & \phi_2(\mathbf{x}_4) \\ \phi_0(\mathbf{x}_5) & \phi_1(\mathbf{x}_5) & \phi_2(\mathbf{x}_5) \end{pmatrix} = \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 1 & 4 & 1 \\ 1 & 3 & 1 \\ 1 & 4 & 2 \end{pmatrix} \end{array}$$

$$R(w^{(\tau)}) = \mathrm{diag}\left(\pi(x_1)\cdot(1-\pi(x_1)), \dots, \pi(x_5)(1-\pi(x_5))\right)$$

$$a = w^T x$$

$$\pi(x)(1-\pi(x)) = \frac{\exp(a)}{1+\exp(a)}\left(1 - \frac{\exp(a)}{1+\exp(a)}\right)$$

$$= \frac{\exp(a)}{1+\exp(a)}\left(\frac{1}{1+\exp(a)}\right)$$

$$= \frac{\exp(a)}{(1+\exp(a))^2} = \frac{\exp(w^T x)}{(1+\exp(w^T x))^2}$$

Made with Goodnotes

Startwert: $w^{(0)} = (0,0,0)^T$

$$\pi(x)\cdot(1-\pi(x)) = \frac{e(0)}{(1+e(0))^2}$$

$$= \frac{1}{2^2} = 1/4$$

$$R(w^{(0)}) = \begin{pmatrix} 1/4 & & & & 0 \\ & 1/4 & & & \\ & & 1/4 & & \\ & & & 1/4 & \\ 0 & & & & 1/4 \end{pmatrix}$$

$$\pi(x) = \frac{\exp(w^T x)}{1+\exp(w^T x)}$$

mit $w^{(0)} = (0,0,0)^T$

$$= \frac{e(0)}{1+e(0)} = 1/2$$

$\Phi \qquad w^{(0)} \qquad R(w^{(0)})^{-1} \qquad \pi(w^{(0)}) \qquad t$

$$z(w^{(0)}) = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 1 & 4 & 1 \\ 1 & 3 & 1 \\ 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \left(\begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} -2 \\ -2 \\ -2 \\ 2 \\ 2 \end{pmatrix}$$

in II $\qquad w^{(1)} = \begin{pmatrix} -20/29 \\ 14/29 \\ -14/29 \end{pmatrix} \approx \begin{pmatrix} -0{,}69 \\ 0{,}48 \\ -0{,}48 \end{pmatrix}$

→ benutzt man um $R(w^{(1)})$

$\pi(w^{(1)})$

→ $z(w^{(1)})$

→ wieder in II einsetzen

$w^{(2)} \rightarrow \begin{pmatrix} -0{,}574 \\ 0{,}61 \\ -0{,}61 \end{pmatrix}$

$$\mathbf{w}^{(3)} \approx \begin{pmatrix} -1{,}046 \\ 0{,}641 \\ -0{,}641 \end{pmatrix}, \ \mathbf{w}^{(4)} \approx \begin{pmatrix} -1{,}05 \\ 0{,}642 \\ -0{,}642 \end{pmatrix}, \ \mathbf{w}^{(5)} \approx \begin{pmatrix} -1{,}05 \\ 0{,}642 \\ -0{,}642 \end{pmatrix}, \ \mathbf{w}^{(6)} \approx \begin{pmatrix} -1{,}05 \\ 0{,}642 \\ -0{,}642 \end{pmatrix}$$

$\Rightarrow \hat{w}_0 = -1{,}05 \quad // \quad \hat{w}_1 = 0{,}642, \ \hat{w}_2 = -0{,}642$

$$\hat{\pi}(x) = \frac{\exp(\hat{w}_0 + \hat{w}_1 x_1 + \hat{w}_2 x_2)}{1 + \exp(\hat{w}_0 + \hat{w}_1 x_1 + \hat{w}_2 x_2)}$$

$$= \frac{\exp(-1{,}05 + 0{,}642 x_1 - 0{,}642 x_2)}{1 + \exp(-1{,}05 + 0{,}642 x_1 - 0{,}642 x_2)}$$

$$\hat{y}^*(x) = \begin{cases} C_1 & \text{für } \hat{\pi}(x) \leq 0{,}5 \\ C_2 & \text{für } \hat{\pi}(x) > 0{,}5 \end{cases}$$

Bayes-Klassifikator

Das ist äquivalent zu

$$\hat{y}^*(x) = \begin{cases} C_1 & \text{für } -1{,}05 + 0{,}642 x_1 - 0{,}642 x_2 \leq 0 \\ C_2 & \text{für } -1{,}05 + 0{,}642 x_1 - 0{,}642 x_2 > 0 \end{cases}$$

für neue Beobachtung $(5,0)$

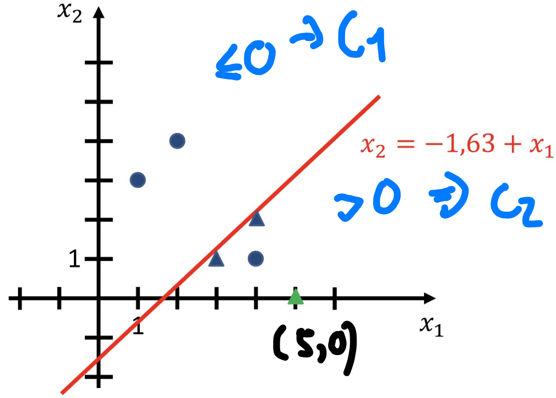
$$-1{,}05 + 0{,}642 \cdot 5 - 0{,}642 \cdot 0 > 0$$

$\rightarrow$ Zuordnung zu Klasse C_2

Für Abbildung der Entscheidungsgrenze:

$$-1{,}05 + 0{,}642 x_1 - 0{,}642 x_2 = 0$$

$$x_2 = -1{,}635514 + x_1$$


x_2
$\leq 0 \to C_1$
$x_2 = -1,63 + x_1$
$> 0 \to C_2$
1
1
x_1
(5,0)