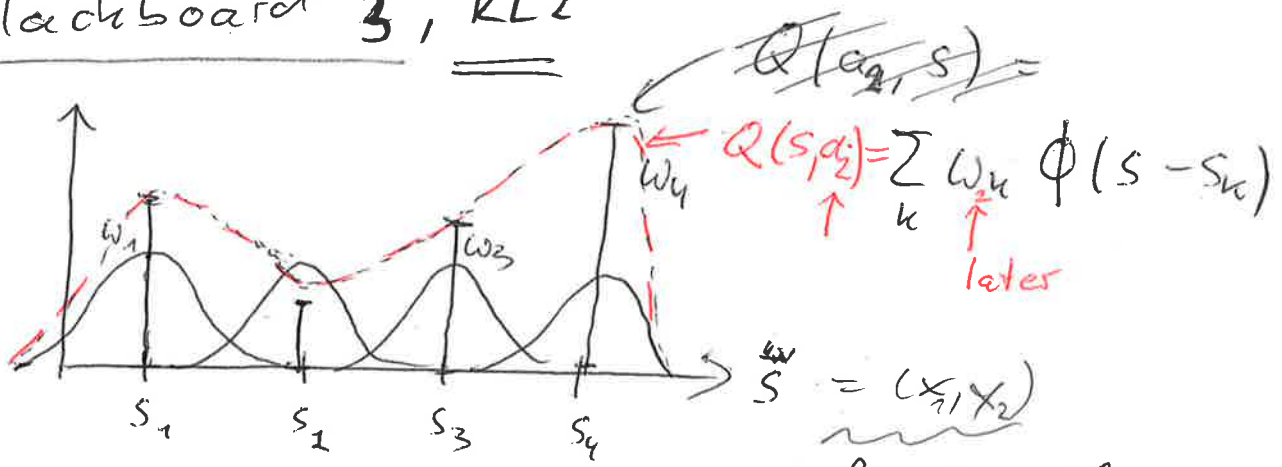
Blackboard 3, RL2



amplitudes

$\omega_1 \quad \omega_2 \quad \omega_3 \quad \omega_4$

$\Rightarrow$ smooth function with few parameters

$Q(s, a_2)$: $\omega_{21}, \omega_{22}, \omega_{23}, \omega_{24}$

general

$Q(s, a \mid \theta)$ — all weights!

$a \in \{a_1, \dots, a_K\}$ K discrete action choices

$s = (x_1, x_2, \dots x_N)$

continuous state space

Blackboard 4 – RL 2: Loss function

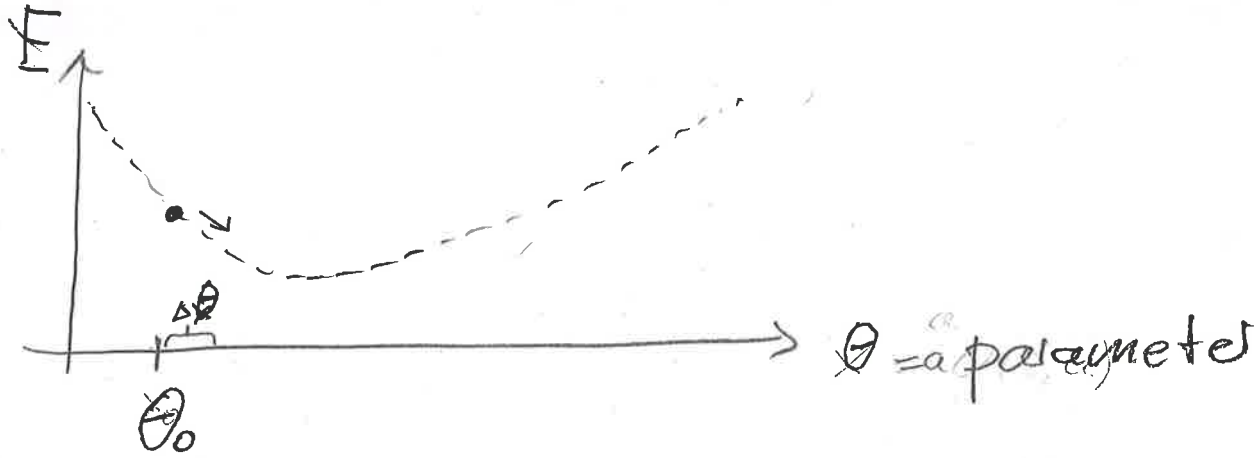
error (loss function)

$$E(\vec{\omega}^{a_1}, \dots) = \frac{1}{2}\left[\underbrace{r + \gamma Q(s', a'|\theta)}_{\text{target}} - \underbrace{Q(s, a|\theta)}_{\text{depends on parameters } \theta}\right]^2$$

$$= \frac{1}{2}[\delta]^2$$

(θ = the weights $\omega_{11}, \omega_{12}, \dots$ $\omega_{21}, \dots$)

minimize error by gradient descent



$$\Delta\theta = -\eta \cdot \frac{\partial E}{\partial \theta} = +\eta \cdot \left[\underbrace{r + \gamma Q(s', a'|\theta)}_{\text{target: fixed}} - Q(s, a|\theta)\right] \frac{\partial Q(s, a|\theta)}{\partial \theta}$$

for basis function network with weights $\vec{\omega}^{a_1}, \vec{\omega}^{a_2}$

$$\Delta \omega_n^{a_1} = -\eta \cdot \frac{\partial E}{\partial \omega_n^{a_1}} = +\eta \left[r + \gamma Q(s', a'|\vec{\omega}^{a'}) - Q(s, a|\vec{\omega}^{a_1})\right] \cdot \frac{\partial Q(s, a|\vec{\omega}^{a_1})}{\partial \omega_n^{a_1}}$$

exercise in class

$$= \underbrace{\delta_{a_1, a_n}}_{\text{Kronecker}} \cdot \phi(s - s_n)$$