Tipps zur Programmierung von Aufgabe 3.2 b) - e):

- Der Gradient ist:

$$\nabla E(\underline{w}) = \begin{pmatrix} \frac{\partial E}{\partial w_1} \\ \frac{\partial E}{\partial w_2} \\ \frac{\partial E}{\partial w_3} \end{pmatrix}$$

  mit:

$$\begin{aligned}
\frac{\partial E}{\partial w_1} &= \frac{2w_1}{\beta} e^{-\frac{w_1^2+w_2^2+w_3^2}{\beta}} + \alpha \sin(w_1)\cos(w_2)\cos(w_3) \\
\frac{\partial E}{\partial w_2} &= \frac{2w_2}{\beta} e^{-\frac{w_1^2+w_2^2+w_3^2}{\beta}} + \alpha \cos(w_1)\sin(w_2)\cos(w_3) \\
\frac{\partial E}{\partial w_3} &= \frac{2w_3}{\beta} e^{-\frac{w_1^2+w_2^2+w_3^2}{\beta}} + \alpha \cos(w_1)\cos(w_2)\sin(w_3)
\end{aligned}$$

- Verschiedene Werte für $\underline{w}_{init}$ erhält man, indem man den Vektor mit zufälligen Werten füllt. Jeder dieser drei w-Werte sollte gleichverteilt zufällig aus dem Bereich [-10, 10] gezogen werden.
- $\Delta\underline{w} = -\eta \cdot \nabla E(\underline{w})$ stellt die Gewichts**änderung** dar. Dies sind **nicht** die neuen Gewichte!
- Das Abbruchkriterium in 3.2c) muss auch den divergierenden Fall berücksichtigen!
- Die Fehlerwerte in den Minima sollten im vorgegebenen Bereich ungefähr zwischen 0 und 0.5 liegen.