

### A1.1

$$\partial E / \partial w(t) = 4 * w(t)$$

$$m(t) = 0.9 * m(t-1) + \partial E / \partial w(t-1) = 0.9 * m(t-1) + 4 * w(t-1)$$

$$w(t) = w(t-1) - 0.1 * m(t)$$

t	m(t)	w(t)
0	0	20
1	$0.9 * 0 + 4 * 20 = 80$	$20 - 0.1 * 80 = 12$
2	$0.9 * 80 + 4 * 12 = 120$	$12 - 0.1 * 120 = 0$

➔  $w(1) = 12, w(2) = 0$

### A1.2

	$\partial E / \partial w(1) < 0$	$\partial E / \partial w(1) > 0$
$\partial E / \partial w(0) < 0$	(+)	(-)
$\partial E / \partial w(0) > 0$	(-)	(+)

### A1.3

- dies liegt daran, dass  $\alpha$  zu groß gewählt wurde. Denn so werden das zuvor berechnete Momentum stärker gewichtet. Da dieses vor dem Überschreiten der 0 recht groß war, ist somit die Schrittweite zu groß und das Minimum wird überschritten.
- Das liegt daran, dass das Momentum stärker steigt als das Gewicht fällt.
- Zu Oszillationen kommt es, wenn  $\alpha$  zu groß gewählt wird, z.B.  $\alpha=1$  (mit entsprechend vielen Iterationen -->  $t=600$ ).
- Mit der Konstellation  $(0.04242, 0.42, 18)^1$  haben wir den minimalen Fehler  $(-0.0015555, 0.001165)$  erreicht.  
([https://milania.de/blog/Introduction\\_to\\_neural\\_network\\_optimizers\\_%5Bpart\\_1%5D\\_%E2%80%93\\_momentum\\_optimization?w1Start=-15&w2Start=20&w1Range=%5B-4.871318018713662%2C4.431007562681689%5D&w2Range=%5B-6.755038948946222%2C2.5472866324491292%5D&error\\_function=3+\\*+w1+%5E+2+%2B+10+\\*+w2+%5E+2&iterations=20&eta=0.04242&alpha=0.42&visible\\_traces=%5B0%2C2%5D#fig:MomentumOptimizer\\_Trajectory](https://milania.de/blog/Introduction_to_neural_network_optimizers_%5Bpart_1%5D_%E2%80%93_momentum_optimization?w1Start=-15&w2Start=20&w1Range=%5B-4.871318018713662%2C4.431007562681689%5D&w2Range=%5B-6.755038948946222%2C2.5472866324491292%5D&error_function=3+*+w1+%5E+2+%2B+10+*+w2+%5E+2&iterations=20&eta=0.04242&alpha=0.42&visible_traces=%5B0%2C2%5D#fig:MomentumOptimizer_Trajectory))