1.

a) *epoch* 1:

$$w1 = -5, w2 = -2, \alpha = 0.4$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-5) = -1$$

$$\nabla f_{w2} = 2.2.(-2) = -8$$

$$w1 = -5 - 0{,}4.(-1) = -4{,}6$$

$$w2 = -2 - 0{,}4.(-8) = 1{,}2$$

epoch 2:

$$w1 = -4{,}6, w2 = 1{,}2, \alpha = 0.4$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-4{,}6) = -0{,}92$$

$$\nabla f_{w2} = 2.2.(1{,}2) = 4{,}8$$

$$w1 = -4{,}6 - 0{,}4.(-0{,}92) = -4{,}232$$

$$w2 = 1{,}2 - 0{,}4.(4{,}8) = -0{,}72$$

2

a) *epoch* 1:

$$w1 = -5, w2 = -2, \alpha = 0.6, v1 = 0, v2 = 0,$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-5) = -1$$

$$v1 = 0{,}5.0 + (1 - 0{,}5).-1 = -0{,}5$$

$$\nabla f_{w2} = 2.2.(-2) = -8$$

$$v2 = 0{,}5.0 + (1 - 0{,}5).-8 = -4$$

$$w1 = -5 - 0{,}6.(-1) = -4{,}4$$

$$w2 = -2 - 0{,}6.(-4) = 0{,}4$$

epoch 1:

$$w1 = -4{,}4, w2 = 0{,}4, \alpha = 0.6, v1 = -0.5, v2 = -4$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-4{,}4) = -0{,}88$$

$$v1 = 0{,}5.(-0{,}5) + (1 - 0{,}5).-0{,}88 = -0{,}69$$

$$\nabla f_{w2} = 2.2.(0{,}4) = 1{,}6$$

$$v2 = 0{,}5.(-4) + (1 - 0{,}5).1{,}6 = -1.2$$

$$w1 = -4{,}4 - 0{,}6.(-0{,}69) = -3{,}986$$

$$w2 = -0{,}4 - 0{,}6.(-1{,}2) = -0{,}32$$

3.

a) *epoch* 1:

$$w1 = -5, w2 = -2, \alpha = 0.3, s1 = 0, s2 = 0, \gamma = 0.9, \epsilon = 10^{-6}$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-5) = -1$$

$$s1 = 0{,}9.0 + (1 - 0{,}9).(-1)^2 = 0{,}1$$

$$\nabla f_{w2} = 2.2.(-2) = -8$$

$$s2 = 0{,}9.0 + (1 - 0{,}9).(-8)^2 = 6{,}4$$

$$w1 = -5 - \frac{0{,}3}{\sqrt{0{,}1 + 10^{-6}}} = -5{,}95$$

$$w2 = -2 - \frac{0{,}3}{\sqrt{6{,}4 + 10^{-6}}} = -2{,}12$$

epoch 2:

$$w1 = -5{,}95, w2 = -2{,}12, \alpha = 0.3, s1 = 0{,}1, s2 = 6{,}4, \gamma = 0.9, \epsilon = 10^{-6}$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-5{,}95) = -1{,}19$$

$$s1 = 0{,}9.0{,}1 + (1 - 0{,}9).(-1{,}19)^2 = 0{,}232$$

$$\nabla f_{w2} = 2.2.(-2{,}12) = -8{,}48$$

$$s2 = 0{,}9.6{,}4 + (1 - 0{,}9).(-8{,}48)^2 = 12{,}95$$

$$w1 = -5{,}95 - \frac{0{,}3}{\sqrt{0{,}232 + 10^{-6}}} = -6{,}57$$

$$w2 = -2{,}12 - \frac{0{,}3}{\sqrt{12{,}95 + 10^{-6}}} = -2{,}03$$

4.

a) *epoch* 1:

$$w1 = -5, w2 = -2, \alpha = 0.2, s1 = 0, s2 = 0, v1 = 0, v2 = 0, \beta1 = 0.9, \beta2 = 0.999, \epsilon = 10^{-6}$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-5) = -1$$

$$v1 = 0{,}9.0 + (1 - 0{,}9).(-1) = -0{,}1$$

$$v1_{corr} = \frac{-0{,}1}{1 - 0{,}9^1} = -1$$

$$s1 = 0{,}999.0 + (1 - 0{,}999).(-1)^2 = 0{,}001$$

$$s1_{corr} = \frac{0.001}{1 - 0{,}999^1} = 1$$

$$\nabla f_{w2} = 2.2.(-2) = -8$$

$$v2 = 0{,}9.0 + (1 - 0{,}9).(-8) = -0{,}8$$

$$v2_{corr} = \frac{-0{,}8}{1 - 0{,}9^1} = -8$$

$$s2 = 0{,}999.0 + (1 - 0{,}999).(-8)^2 = 0.064$$

$$s2_{corr} = \frac{0.064}{1 - 0{,}999^1} = 64$$

$$w1 = -5 - \frac{0{,}2.(-1)}{\sqrt{1 + 10^{-6}}} = -4{,}8$$

$$w2 = -2 - \frac{0{,}2.(-8)}{\sqrt{64 + 10^{-6}}} = -1{,}8$$

epoch 2:

$$w1 = -4{,}8, w2 = -1{,}8, \alpha = 0.2, s1 = 0{,}001, s2 = 0{,}064, v1 = -0{,}1, v2 = -0{,}8, \beta1 = 0.9, \beta2 = 0.999, \epsilon = 10^{-6}$$

$$f(w1, w2) = 0{,}1w1^2 + 2w2^2$$

$$\nabla f_{w1} = 2.0{,}1.(-4{,}7) = -0{,}94$$

$$v1 = 0{,}9.(-0{,}1) + (1 - 0{,}9).(-0{,}94) = -0{,}184$$

$$v1_{corr} = \frac{-0,184}{1 - 0,9^2} = -0,97$$

$$s1 = 0,999.\,(0,001) + (1 - 0,999).\,(-0,94)^2 = 0,002$$

$$s1_{corr} = \frac{0,002}{1 - 0,999^2} = 0,942$$

$$\nabla f_{w2} = 2.2.\,(-1,7) = -6,8$$

$$v2 = 0,9.\,(-0,8) + (1 - 0,9).\,(-6.8) = -1,4$$

$$v2_{corr} = \frac{-1,4}{1 - 0,9^2} = -7,37$$

$$s2 = 0,999.0,064 + (1 - 0,999).\,(-6,8)^2 = 0.11$$

$$s2_{corr} = \frac{0.11}{1 - 0,999^2} = 55,03$$

$$w1 = -4,8 - \frac{0,2.\,(-0,97)}{\sqrt{0,942} + 10^{-6}} = -4,6$$

$$w2 = -1,8 - \frac{0,2.\,(-7,37)}{\sqrt{55,03} + 10^{-6}} = -1,6$$