Solutions

2.1

- $w_1 = 4 - 0.2(1) = 3.8$; $w_2 = 3.8 - 0.2(0.8) = 3.64$ ✓.

2.2

- $|1 - 2\eta| < 1 \iff 0 < \eta < 1$ ✓.

2.3

- A ravine: 100x curvature mismatch between the axes ✓.

3.1

- Update: $w_{t+1} = w_t - 0.5(w_t + 2) = 0.5w_t - 1$ ✓.
- Iterates: $1 \to -0.5 \to -1.25 \to -1.625 \to -1.8125$ ✓.
- Limit -2 (the minimum); errors shrink monotonically, no overshoot ✓.
- With $\eta = 0.75$: $w_{t+1} = w_t - 1.5(w_t + 2) = -0.5w_t - 3$: $1 \to -3.5 \to -1.25 \to -2.375 \to -1.8125$ – alternating sign of the error, damped oscillation ✓.

3.2

- Curvatures: 2 and 200 (the loss's second derivatives) ✓.
- Shared-rate stability: $\eta < 2/200 = 0.01$ (the tighter axis rules) ✓.
- Halving the error per coordinate: the slow axis contracts at $|1 - 2\eta| = 0.98$ per step – about 34 steps to halve ✓.
- Standardising the second feature equalises curvature at 2: the bound becomes $\eta < 1$ and one rate serves both axes ✓.

exercises.lol exercises.com.cn