

**7.2** Consider the update equation for stochastic gradient descent (Equation (7.15)).

Write down the update when we use a mini-batch size of one.

**Solution.**

$$\theta_{i+1} = \theta_i - \gamma_i (\nabla L_n(\theta_i))^T$$

where  $n$  is a random number between 1 and  $N$