1. (c) We solve

$$0.008 \leq 0.1^2 \left(1 - \frac{8+1}{N}\right)$$

$$N \geq \frac{1+8}{1 - 0.008/0.1^2} = 45$$

2. (d) This is a hyperbola, and because we want -1 in $\mathcal{X}$ space as $x_1 \to \infty$, it follows that $w_1 < 0$.

3. (c) 14th order space in the non-linear space, so our VC dimension is 14.

4. (e) Just take the derivative.

5. (d) We get 10 iterations to reach $E_{\text{in}} < 10^{-14}$.

6. (e) We get (0.0447, 0.024)

7. (a) Using the two-step method sucks because you're not going in the optimal direction.

8. (d) I got 0.108

9. (a) I got 340 average

10. (e) The PLA takes the update
$$\mathbf{w} + y_n \mathbf{x}_n \to \mathbf{w}$$

Thus when performing SGD on point n, it follows that

$$\nabla e_n(\mathbf{w}) = -y_n \mathbf{x}_n$$

so

$$e_n(\mathbf{w}) = -y_n \mathbf{w}^T \mathbf{x}_n$$

However, we only perform this weight update if point n has an error. Thus we have

$$e_n(\mathbf{w}) = -\min(0, y_n \mathbf{w}^T \mathbf{x}_n).$$