Recitation 13 Solutions

1. (a) $h_w(\mathbf{x}_i) = \mathbf{w} \cdot \mathbf{x}_i$

   (b) $J_n(\mathbf{w}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mathbf{w} \cdot \mathbf{x}_i)^2$.

   (The factor of $\frac{1}{n}$ is optional; it's constant, so it doesn't change the minimizing $\mathbf{w}$.)

   (c) Take the derivative of $J_n(\mathbf{w})$ with respect to each element of $\mathbf{w}$ (w_0 and w_1) and set it to 0.

   $\frac{\partial J_n}{\partial w_j} = -2\sum_{i=1}^{n}(y_i - (w_0 x_{i0} + w_1 x_{i1}))x_{ij} = 0$ for all j.

   (Note that x_{i0} is always 1.)

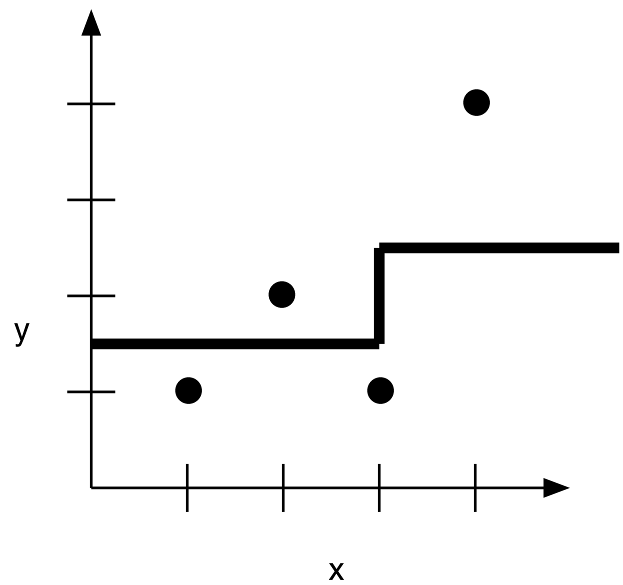
   (d) Yes. The formula is at the bottom of the first page of the handout.

   (e) Yes, we're just using an input vector that happens to be $\langle 1, x, x^2 \rangle$ instead of $\langle 1, x \rangle$.

2. See Figure 18.14 on page 722 of AIMA third edition. If I have time maybe I'll scan it for you.

3.

1. 2-nearest-neighbor (equally weighted averaging)



2. regression trees (with leaf size 1)