Exercise: Fitting an SVM classifier by hand

(Source: Jaakkola.) Consider a dataset with 2 points in 1d: $(x_1 = 0, y_1 = -1)$ and $(x_2 = \sqrt{2}, y_2 = 1)$. Consider mapping each point to 3d using the feature vector $\boldsymbol{\phi}(x) = [1, \sqrt{2}x, x^2]^T$. (This is equivalent to using a second order polynomial kernel.) The max margin classifier has the form

$$\min ||\mathbf{w}||^2 \quad \text{s.t.} \tag{1}$$

$$y_1(\mathbf{w}^T \phi(\mathbf{x}_1) + w_0) \geq 1 \tag{2}$$

$$y_2(\mathbf{w}^T \phi(\mathbf{x}_2) + w_0) \geq 1 \tag{3}$$

1. Write down a vector that is parallel to the optimal vector $\mathbf{w}$. Hint: recall from Figure 7.8 (12Apr10 version) that $\mathbf{w}$ is perpendicular to the decision boundary between the two points in the 3d feature space.
2. What is the value of the margin that is achieved by this $\mathbf{w}$? Hint: recall that the margin is the distance from each support vector to the decision boundary. Hint 2: think about the geometry of 2 points in space, with a line separating one from the other.
3. Solve for $\mathbf{w}$, using the fact the margin is equal to $1/||\mathbf{w}||$.
4. Solve for w_0 using your value for $\mathbf{w}$ and Equations 1 to 3. Hint: the points will be on the decision boundary, so the inequalities will be tight.
5. Write down the form of the discriminant function $f(x) = w_0 + \mathbf{w}^T \boldsymbol{\phi}(x)$ as an explicit function of x.