Answer 2.2: Margin

a. Example for ε margin halfspace:

It happens when an arbitrary working hyperplane is formed biased towards the last mistaken points.

Simplest Example:

$$(x_i, y_i) = \begin{cases} ((2,2),1) & i = 2n+1 \\ ((-0.5,-0.5),-1) & i = 2n+2 \end{cases}, \forall n \in \mathbb{Z}$$

Proof:

Iteration 1: Since $y_1(\langle w_0, x_1\rangle + b_0) = 0 \leq 0, \Rightarrow \quad w_1 = w_0 + y_i x_i = (2,2)$ and $b_1 = b_0 + y_1 = 1$

Iteration 2: Since $y_2(\langle w_1, x_2\rangle + b_1) = -1 * (\langle(-0.5,-0.5),(2,2)\rangle + 1) = 1 > 0, \Rightarrow$ Do not update

Iteration 3: Since $y_3(\langle w_2, x_3\rangle + b_2) = 1 * (\langle(2,2),(2,2)\rangle + 1) = 9 > 0, \Rightarrow$ Do not update

Iteration 2i: As stated in Iteration 2

Iteration 3i: As stated in Iteration 3

Hence, it will never update after the first update.

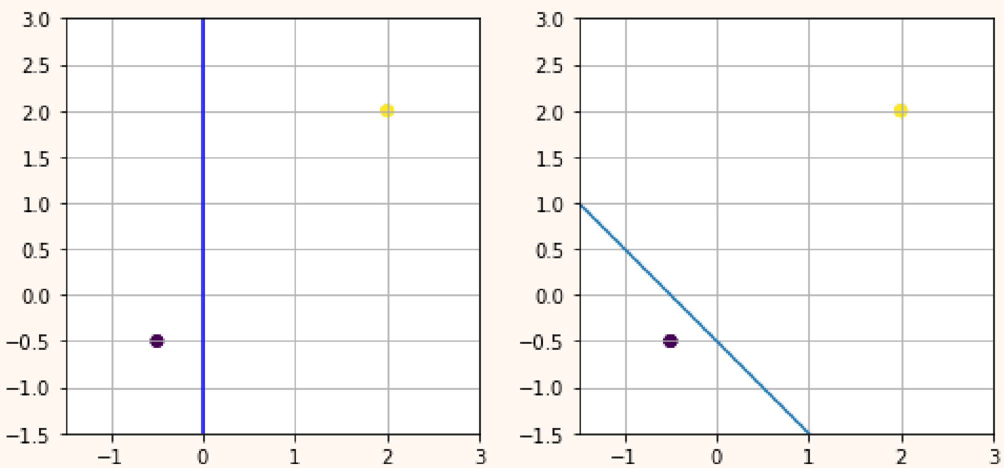
The converged hyperplane is defined by $w = (2,2), b = 1$

Hence, we may compute the margin: $\gamma = 0.5/\sqrt{2} = 0.3535 < 1/2$

And, the max geometric margin: $\hat{\gamma} = \|(2,2) - (-0.5,-0.5)\| = 3.5355 > 1$

Hence, in this case, the margin obtained is smaller than the max possible geometric margin.

Initial Hyperplane $\rightarrow$ Final Hyperplane



b. Example for a maximum margin halfspace:
It happens when $\gamma = R = \max_i ||a_i||_2$,
hence, all the data points are concentrated on two symmetrical points with opposing label.
It would also make one mistake $\lim_{\gamma \to R} (\frac{R}{\gamma})^2 = 1$
Example: $(x_i, y_i) = (((-1)^i, 0), (-1)^i)$, $\gamma = R = \max_i ||a_i||_2 = 1$