50.6 Hard Margin Support Vector Machine; Version II

Since $\delta > 0$ (otherwise the data would not be separable into two disjoint sets), we can divide the affine constraints by δ to obtain

$$\begin{aligned} w'^{\top} u_i - b' &\geq 1 && i = 1, \ldots, p \\ -w'^{\top} v_j + b' &\geq 1 && j = 1, \ldots, q, \end{aligned}$$

except that now, w' is not necessarily a unit vector. To obtain the distances to the hyperplane H, we need to divide by $\|w'\|$ and then we have

$$\begin{aligned} \frac{w'^{\top} u_i - b'}{\|w'\|} &\geq \frac{1}{\|w'\|} && i = 1, \ldots, p \\ \frac{-w'^{\top} v_j + b'}{\|w'\|} &\geq \frac{1}{\|w'\|} && j = 1, \ldots, q, \end{aligned}$$

which means that the shortest distance from the data points to the hyperplane is $1/\|w'\|$. Therefore, we wish to maximize $1/\|w'\|$, that is, to minimize $\|w'\|$, so we obtain the following optimization Problem (SVM_{h2}):

Hard margin SVM (SVM_{h2}):

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\|w\|^2 \\ &\text{subject to} \\ &\qquad w^{\top} u_i - b \geq 1 \qquad i = 1, \ldots, p \\ &\qquad -w^{\top} v_j + b \geq 1 \qquad j = 1, \ldots, q. \end{aligned}$$

The objective function $J(w) = 1/2\|w\|^2$ is convex, so Proposition 50.7 applies and gives us a necessary and sufficient condition for having a minimum in terms of the KKT conditions. First observe that the trivial solution $w = 0$ is impossible, because the blue constraints would be

$$-b \geq 1,$$

that is $b \leq -1$, and the red constraints would be

$$b \geq 1,$$

but these are contradictory. **Our goal is to find w and b, and optionally, δ.** We proceed in four steps first demonstrated on the following example.

Suppose that $p = q = n = 2$, so that we have two blue points

$$u_1^{\top} = (u_{11}, u_{12}) \qquad u_2^{\top} = (u_{21}, u_{22}),$$