

It will also be useful to understand how points are classified in terms of  $\epsilon_i$  (or  $\xi_j$ ).

- (1) If  $\epsilon_i > 0$ , then by complementary slackness  $\lambda_i = K_s$ , so the  $i$ th equation is active and by (2) above,

$$w^\top u_i - b - \eta = -\epsilon_i.$$

Since  $\epsilon_i > 0$ , the point  $u_i$  is strictly within the open half space bounded by the blue margin hyperplane  $H_{w,b+\eta}$  and containing the separating hyperplane  $H_{w,b}$  (excluding the blue hyperplane  $H_{w,b+\eta}$ ); if  $\epsilon_i \leq \eta$ , then  $u_i$  is classified correctly, and if  $\epsilon_i > \eta$ , then  $u_i$  is misclassified.

Similarly, if  $\xi_j > 0$ , then  $v_j$  is strictly within the open half space bounded by the red margin hyperplane  $H_{w,b-\eta}$  and containing the separating hyperplane  $H_{w,b}$  (excluding the red hyperplane  $H_{w,b-\eta}$ ); if  $\xi_j \leq \eta$ , then  $v_j$  is classified correctly, and if  $\xi_j > \eta$ , then  $v_j$  is misclassified.

- (2) If  $\epsilon_i = 0$ , then the point  $u_i$  is correctly classified. If  $\lambda_i = 0$ , then by (3) above,  $u_i$  is in the closed half space on the blue side bounded by the blue margin hyperplane  $H_{w,b+\eta}$ . If  $\lambda_i > 0$ , then by (1) and (2) above, the point  $u_i$  is on the blue margin.

Similarly, if  $\xi_j = 0$ , then the point  $v_j$  is correctly classified. If  $\mu_j = 0$ , then  $v_j$  is in the closed half space on the red side bounded by the red margin hyperplane  $H_{w,b-\eta}$ , and if  $\mu_j > 0$ , then the point  $v_j$  is on the red margin.

Also observe that if  $\lambda_i > 0$ , then  $u_i$  is in the closed half space bounded by the blue hyperplane  $H_{w,b+\eta}$  and containing the separating hyperplane  $H_{w,b}$  (including the blue hyperplane  $H_{w,b+\eta}$ ).

Similarly, if  $\mu_j > 0$ , then  $v_j$  is in the closed half space bounded by the red hyperplane  $H_{w,b-\eta}$  and containing the separating hyperplane  $H_{w,b}$  (including the red hyperplane  $H_{w,b-\eta}$ ).

**Definition 54.3.** Vectors  $u_i$  such that  $\lambda_i > 0$  and vectors  $v_j$  such that  $\mu_j > 0$  are said to *have margin at most  $\delta$* . The sets of indices associated with these vectors are denoted by

$$I_{\lambda>0} = \{i \in \{1, \dots, p\} \mid \lambda_i > 0\}$$

$$I_{\mu>0} = \{j \in \{1, \dots, q\} \mid \mu_j > 0\}.$$

We denote their cardinalities by  $p_m = |I_{\lambda>0}|$  and  $q_m = |I_{\mu>0}|$ .

Vectors  $u_i$  such that  $\epsilon_i > 0$  and vectors  $v_j$  such that  $\xi_j > 0$  are said to *strictly fail the margin*. The corresponding sets of indices are denoted by

$$E_\lambda = \{i \in \{1, \dots, p\} \mid \epsilon_i > 0\}$$

$$E_\mu = \{j \in \{1, \dots, q\} \mid \xi_j > 0\}.$$

We write  $p_{sf} = |E_\lambda|$  and  $q_{sf} = |E_\mu|$ .