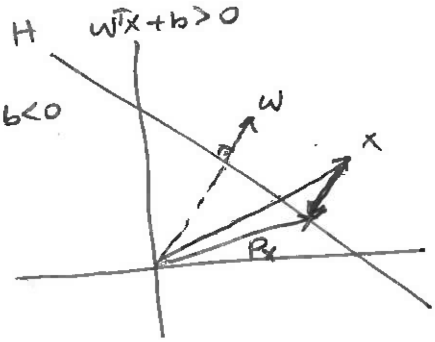
$$H(w,b) = \{ x : w^T x + b = 0 \} \qquad \|w\| = 1$$

* Distance from point x to $H(w,b)$. $\quad d(x,H) = |w^T x + b|$



proof. $\quad P_x \triangleq x - (w^T x + b) w$

$$w^T p_x + b = w^T x - w^T (w^T x + b) w + b$$
$$= (w^T x + b) - \overset{1}{\cancel{\|w\|^2}} (w^T x + b)$$
$$= 0$$

$$\| x - p_x \| = |w^T x + b|$$

$(x_1, y_1), \ldots, (x_m, y_m) \qquad y_m \in \{-1, +1\}$

Hard-SVM: $\quad \underset{(w,b):\|w\|=1}{\arg\max} \ \min_{i=1,\ldots,m} |w^T x_i + b| \quad$ s.t. $\ y_i (w^T x_i + b) > 0$

$$(*) \ (w^*, b^*) = \underset{(w,b):\|w\|=1}{\arg\max} \ \min_{i\in[m]} \left(w^T x_i + b \right) y_i$$

$$(**) \ (w_0, b_0) = \underset{(w,b)}{\arg\min} \|w\|^2 \quad \text{s.t.} \quad y_i (w^T x_i + b) \geq 1$$

proof: ~~Let~~ (w^*, b^*) ~~be~~ Let $\gamma^* = \min_{i\in[m]} y_i \left((w^*)^T x_i + b^* \right)$

$$y_i \left((w^*)^T x_i + b^* \right) \geq \gamma^* \quad , \forall i$$

$$\Leftrightarrow \quad y_i \left(\Big(\underbrace{\frac{w^*}{\gamma^*}}_{w'} \Big)^T x_i + \underbrace{\frac{b^*}{\gamma^*}}_{b'} \right) \geq 1 \quad , \forall i$$

$$\|w_0\| \leq \left\| \frac{w^*}{\gamma^*} \right\| = \frac{1}{\gamma^*}$$

$$\underbrace{\frac{1}{\|w_0\|} \cdot y_i (w_0^T x_i + b_0)}_{y_i \cdot \left(\left(\frac{w_0}{\|w_0\|} \right)^T x_i + \frac{b_0}{\|w_0\|} \right)} \geq \frac{1}{\|w_0\|} \geq \gamma^*$$

Hard margin SVM:

$$\min_{w,b} \frac{1}{2}\|w\|^2$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1, \ \forall i$$

Lagrange Duality:

$$\min_{w} f(w)$$

$$\text{s.t. } g_i(w) \leq 0 \qquad i=1,\dots,k$$

$$h_i(w) = 0 \qquad i=1,\dots,l$$

generalized Lagrangian:

$$\mathcal{L}(w,\alpha,\beta) = f(w) + \sum_{i=1}^{k} \alpha_i g_i(w) + \sum_{i=1}^{l} \beta_i h_i(w)$$

α_i, β_i: Lagrange multipliers

$$\theta_p(w) = \max_{\alpha,\beta:\alpha_i \geq 0} \mathcal{L}(w,\alpha,\beta)$$ "primal"

$$p^* = \min_{w} \theta_p(w) = \min_{w} \max_{\alpha,\beta:\alpha_i \geq 0} \mathcal{L}(w,\alpha,\beta)$$ $\rightarrow$ original problem

$$\theta_D(\alpha,\beta) = \min_{w} \mathcal{L}(w,\alpha,\beta)$$ "dual"

$$d^* = \max_{\alpha,\beta:\alpha_i \geq 0} \theta_D(\alpha,\beta) = \max_{\alpha,\beta:\alpha_i \geq 0} \min_{w} \mathcal{L}(w,\alpha,\beta)$$

$$\boxed{d^* \leq p^*}$$

α, β

1	2	4	1	1
2	3	1	2	1
0	2	1	3	1
2	3	1	0	1
2	3	4	3	

w

$1 = d^*$

$2 = p^*$

Suppose f and g_i's are convex, and h_i's are affine, and g_i's are strictly feasible.

Then, there must exist w^*, α^*, β^* s.t. w^* is the solution of the primal problem, α^*, β^* are solutions of the dual problem, $p^* = d^* = \mathcal{L}(w^*, \alpha^*, \beta^*)$

Moreover, w^*, α^*, β^* satisfy Karush-Kuhn-Tucker (KKT) conditions:

$$\frac{\partial}{\partial w_i} \mathcal{L}(w^*, \alpha^*, \beta^*) = 0 \qquad , \forall i=1,\dots,n$$

$$\frac{\partial}{\partial \beta_i} \mathcal{L}(w^*, \alpha^*, \beta^*) = 0 \qquad \forall i=1,\dots,\ell$$

$$\alpha_i^* \cdot g_i(w^*) = 0 \qquad \forall i=1,\dots,k.$$

$$g_i(w^*) \leq 0 \qquad \forall i$$

$$\alpha_i^* \geq 0 \qquad \forall i$$

If w^*, α^*, β^* satisfy KKT conditions, then it's also a solution of the primal and dual problems.

$$\min_{w,b} \frac{1}{2}\|w\|^2$$

$$\text{s.t.} \quad \underbrace{y_i(w^T x_i + b) \geq 1}_{-y_i(w^T x_i + b) + 1 \leq 0} \qquad \forall i=1,\dots,m$$

$$\mathcal{L}(w,b,\alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{m} \alpha_i \left[y_i(w^T x_i + b) - 1 \right]$$

$$\nabla_w \mathcal{L} = w - \sum_{i=1}^{m} \alpha_i y_i x_i = 0 \implies w^* = \sum_{i=1}^{m} \alpha_i y_i x_i$$

$$\frac{\partial}{\partial b}\mathcal{L} = \sum_{i=1}^{m} \alpha_i y_i = 0$$

$$\alpha_i \cdot \left[-y_i(w^T x_i + b) + 1 \right] = 0$$