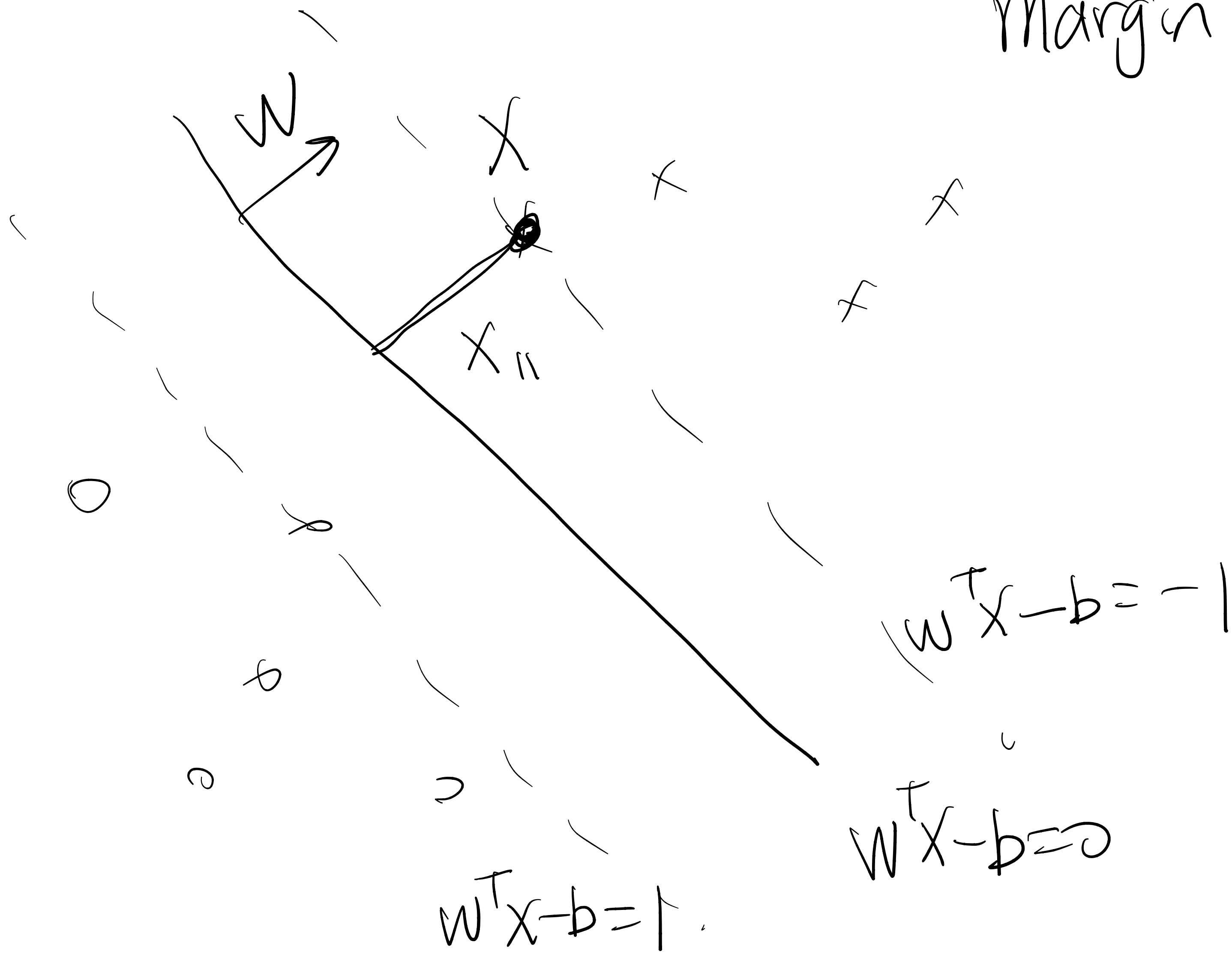
Margin $\|x_{\parallel}\| = \frac{1}{\|w\|}$



Large Margin Classifier.

$$\min \frac{1}{2}\|w\|^2 \quad \text{s.t. } y_i(w^T x_i - b) \geq 1$$

$$\text{for all } i \text{ in } \{1, \cdots, N\}.$$

Optimization w/ constraints.

L(x) with a constraint g(x)

$$\max_{x} L(x) \quad \text{s.t.} \quad g(x) \geq 0.$$

$$\boxed{\lambda \geq 0}$$

$$\Rightarrow \quad L'(x) = L(x) + \lambda g(x).$$

$$\frac{dL'(x)}{dx} = \frac{dL}{dx} + \lambda \frac{dg}{dx} = 0.$$

Karush-Kuhn-Tucker (KKT) condition.

$$\left.\frac{dL}{dx}\right|_{x=x^*} = 0 \qquad \boxed{\lambda = 0} \qquad g(x^*) \geq 0$$

x^* : zero-gradient

$$\left.\frac{dL}{dx}\right|_{x=x'} = 0 \qquad \lambda \geq 0 \qquad \boxed{g(x^*) = 0}$$

: boundary case.

$$g(x') \not\geq 0$$

* Opt. w/ multiple constraints.

$$\min_{X} L(X) = x_1^2 + x_2^2 \qquad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{s.t.} \begin{cases} g_1(X) = x_2 - x_1 - 1 \geq 0 \\ g_2(X) = x_2 + x_1 - 1 \geq 0 \end{cases}$$

$$L'(X) = L(X) - \lambda_1 g_1(X) - \lambda_2 g_2(X)$$

$$L'(X) = x_1^2 + x_2^2 - \lambda_1 (x_2 - x_1 - 1) - \lambda_2 (x_2 + x_1 - 1)$$

$$\frac{\partial L'}{\partial x_1} = 2x_1 + \lambda_1 - \lambda_2 = 0 \qquad \text{—(1)}$$

$$\frac{\partial L'}{\partial x_2} = 2x_2 - \lambda_1 - \lambda_2 = 0 \qquad \text{—(2)}$$

KKT condition

x^* : in the feasible region.

(1) $\lambda_1 = \lambda_2 = 0$.

(2) $\lambda_1 = 0$ $\quad g_2(x) = 0$.

(3) $\lambda_2 = 0$ $\quad g_1(x) = 0$

(4) $g_1(x) = 0$, $\quad g_2(x) = 0$.

~~(1)~~ $\lambda_1=0,\ \lambda_2=0$ 时 $x_1=0,\ x_2=0,\ g_1(0,0)=-1 \not\geq 0$

~~(2)~~ $\lambda_1=0,\ g_2(x)=0$ 时 $x_2+x_1-1=0$

$x_1=x_2=\frac{1}{2}$

$g_1\left(\frac{1}{2},\frac{1}{2}\right)=-1 \not\geq 0$

$$L'(x)=x_1^2+x_2^2-\lambda_1(x_2-x_1-1)-\lambda_2(x_2+x_1-1)$$

$$\frac{\partial L'}{\partial x_1}=2x_1+\lambda_1-\lambda_2=0 \quad —①$$

$$\frac{\partial L'}{\partial x_2}=2x_2-\lambda_1-\lambda_2=0 \quad —②$$

(3) $\lambda_2 = 0$ $(g_2(x) \geq 0)$

$x_2 - x_1 - 1 = 0$ $\Rightarrow$ $x_1 = -\frac{1}{2}, \; x_2 = \frac{1}{2}$

$g_2(-\frac{1}{2}, \frac{1}{2}) = -1 \ngeq 0$

(4) $\begin{cases} x_2 - x_1 - 1 = 0 \\ x_2 + x_1 - 1 = 0 \end{cases}$ $\Rightarrow$ $\boxed{x_2 = 1, \; x_1 = 0}$

$\min \frac{1}{2}\|w\|^2$ s.t. $y_i(w^T x_i - b) \geq 1 \quad i \in \{1, \cdots, N\}$

$$\Rightarrow \boxed{L(w,b) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \{y_i(w^T x_i - b) - 1\} \qquad \alpha_i \geq 0}$$

① $\frac{\partial L}{\partial w} = 0 \Rightarrow w - \sum \alpha_i y_i x_i = 0 \Rightarrow \boxed{w = \sum \alpha_i y_i x_i}$

② $\frac{\partial L}{\partial b} = 0 \Rightarrow \boxed{\sum \alpha_i y_i = 0}$

$\Rightarrow$ plug these eq's into the $L(w,b)$.

① $\|w\|^2 = w^T w = \left(\sum_i \alpha_i y_i x_i^T\right)\left(\sum_i \alpha_i y_i x_i\right)$

$$= \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j .$$

② $\sum \alpha_i y_i w^T x_i = \sum_i \alpha_i y_i \boxed{\sum_j \alpha_j y_j x_j^T} x_i$

$$= \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j .$$

③ $\sum \alpha_i y_i b = b \sum \alpha_i y_i = 0.$

$$L(w, b) = \boxed{\frac{1}{2}\|w\|^2 - \sum \alpha_i \left(y_i (w^T x_i - b) - 1 \right)}$$

$$= \frac{1}{2} \sum\sum \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum\sum \alpha_i \alpha_j y_i y_j x_i^T x_j + \sum \alpha_i$$

$$\boxed{= \sum \alpha_i - \frac{1}{2} \sum\sum \alpha_i \alpha_j y_i y_j x_i^T x_j}$$

$$\begin{cases} a_i \geq 0 \\ w = \sum \alpha_i y_i x_i \\ \sum \alpha_i y_i = 0 \end{cases}$$

Primal

$\min L(w, b)$

$w^T x' - b \gtrless 0$

$D+1$ params

Dual

$\max L(\alpha_i)$

$w^T x' - b$

$= \left(\sum \alpha_i y_i x_i^T \right) x' - b$

N params