Midterm Review

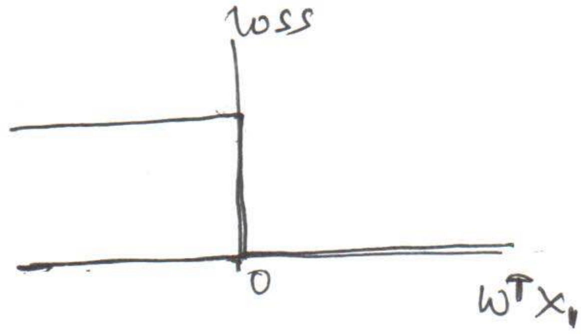
if $\text{sign}\{w^T x\} \cancel{\geq 0} = +1 \quad y = +1$

$= -1 \quad y = -1$

If we have only one training instance x_1, y_1

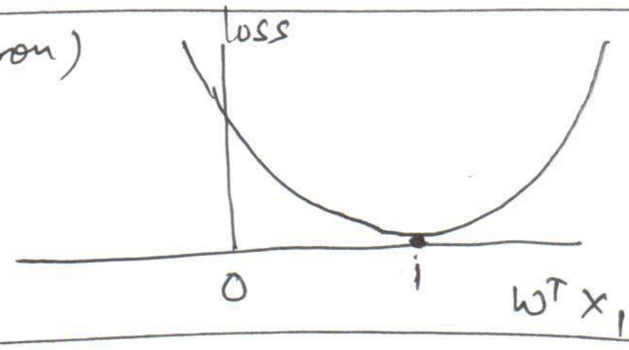
0-1 loss

Assume $y_1 = +1$



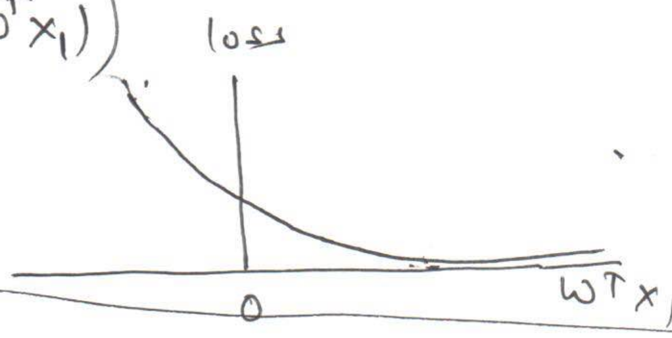
Squared loss. (Perceptron)

$(y_1 - w^T x_i)^2$



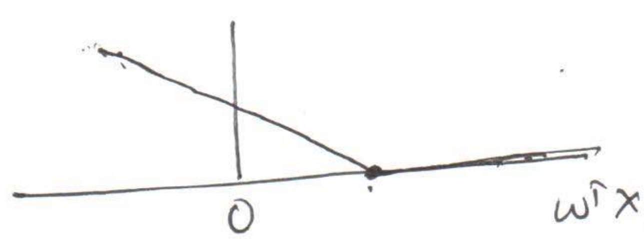
Log-loss (Logistic Reg)

$\log(1 + \exp(-y_1 w^T x_1))$



Hinge-loss (SVM)

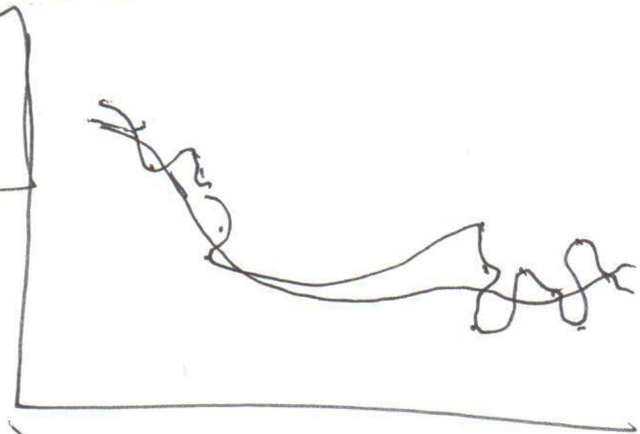
$\max(0, 1 - y_1 w^T x_1)$



$$\frac{1}{2} \sum (y_i - w^T x_i)^2 + \lambda \|w\|_2^2$$

$$\boxed{W = (\lambda I_D + X^T X)^{-1} X^T y}$$

identity matrix.



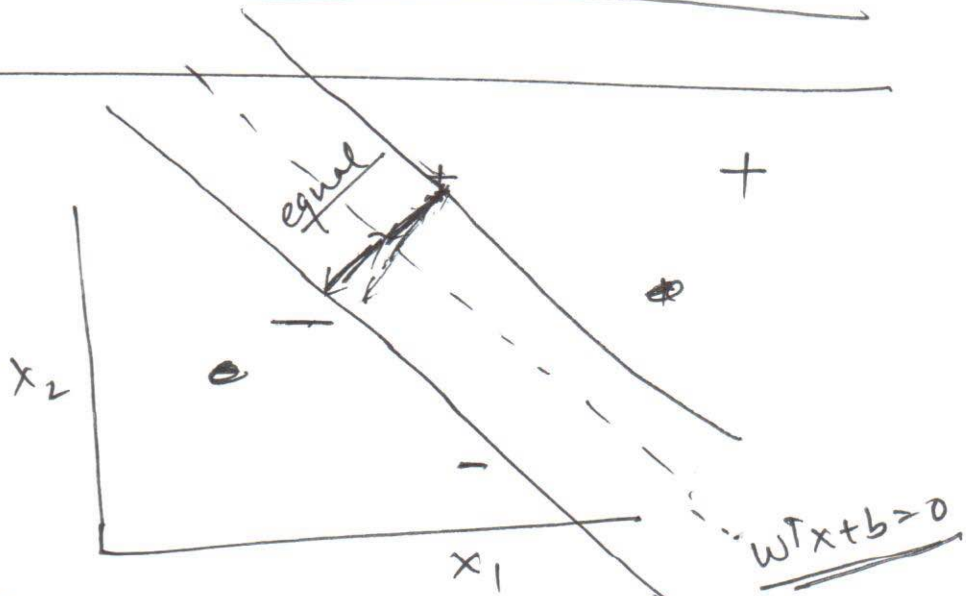
$\cancel{y = mx + c}$

$\boxed{x_2 = m x_1 + c}$

$w_1 x_1 + w_2 x_2 + b = 0$

$2w_1 x_1 + 2w_2 x_2 + 2b = 0$

$$-\frac{w_1}{w_2} = m$$



$$\boxed{\begin{aligned} w_1 x_1 + w_2 x_2 + b &= +1 \\ w_2 x_2 + w_2 x_2 + b &= -1 \end{aligned}}$$

Data

$x_1 \longrightarrow$ $\phi(x_1)$ $\rightarrow [\phi_1(x_1) \quad \phi_2(x_1) \dots \phi_p(x_1)]$

x_2 $\phi(x_2)$

x_3 $\phi(x_3)$

$\vdots$

x_N $\boxed{W^T \phi(x_i)}$

$\boxed{(x_i^T x_j)}$ $\boxed{\phi(x_i)^T \phi(x_j)}$

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

Kernel Methods $\Rightarrow$ No explicit $\phi(\ )$

---

linear regression $\longrightarrow$ Kernel ~~t~~ Regression

"Kernel Trick"

SVM

How to convert a linear model $(w^T x)$ to non-linear model.

Basis fn. expansion

$x_1 \rightarrow \Phi \rightarrow \phi_1, \phi_2, \ldots, \phi_p$

$\phi(x_1)$ [p]

$\phi(x_2)$

x_2

$\boxed{w^T \phi(x)}$

If, for all training examples x_i always occurs as $x_i^T x_j$ $\longrightarrow$ Replace $x_i^T x_j$ with $k(x_i, x_j)$

Kernel Methods

For any kernel fn. $k(x_i, x_j)$ there should be a $\Phi(x_i)$ & $\Phi(x_j)$ such that $k(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$

Radial Basis function Kernel (RBF) $\longrightarrow$ Gaussian kernel

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|_2^2}{\sigma}\right)$$

$x_i, x_j \in \mathbb{R}^d$

example: $d = 2$ $\quad x_i = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ $\quad x_j = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\sigma = 1$

$$k(x_i, x_j) = \exp\left(-\frac{\left\|\begin{matrix} 2 \\ 2 \end{matrix}\right\|_2^2}{1}\right) = \exp\left(-\frac{(2^2+2^2)}{1}\right)$$

$$= \exp(-8)$$

Assume $\sigma = 1$, $d = 1$

$$k(x_i, x_j) = \exp\left(-(x_i - x_j)^2\right)$$

$x_i, x_j \in \mathbb{R}$

$$= \exp\left(-x_i^2 - x_j^2 + 2x_i x_j\right)$$

$$= \exp(-x_i^2)\exp(-x_j^2)\underline{\exp(2x_i x_j)}$$

Maclaurin Series Expansion.

$$\boxed{\exp(y) = \sum_{k=0}^{\infty} \frac{y^k}{k!}}$$

$$\underline{k(x_i, x_j)} = \exp(-x_i^2)\exp(-x_j^2)\sum_{k=0}^{\infty} \frac{2^k x_i^k x_j^k}{k!}$$

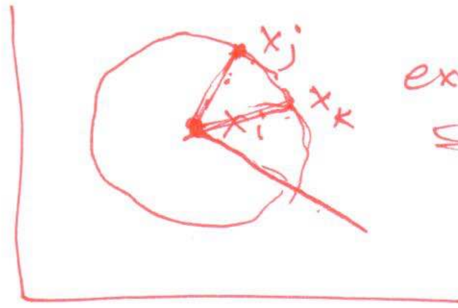
$$= \exp(-x_i^2)\exp(-x_j^2)\sum_{k=0}^{\infty}\left(\frac{2^{k/2} x_i^k}{\sqrt{k!}}\right)\left(\frac{2^{k/2} x_j^k}{\sqrt{k!}}\right)$$

$$= \sum_{k=0}^{\infty}\left[\underbrace{\left(\frac{\exp(-x_i)^2 \cdot 2^{k/2} x_i^k}{\sqrt{k!}}\right)}_{\alpha_{ik}}\underbrace{\left(\frac{\exp(-x_j)^2\, 2^{k/2} x_j^k}{\sqrt{k!}}\right)}_{\alpha_{jk}}\right]$$

$$\sum_{k=0}^{\infty} \alpha_{ik}\alpha_{jk} = \begin{bmatrix}\alpha_{i0}\\ \alpha_{i1}\\ \vdots\\ \alpha_{i\infty}\end{bmatrix} \begin{bmatrix}\alpha_{j0}\\ \alpha_{j1}\\ \vdots\\ \alpha_{j\infty}\end{bmatrix}$$

RBF $\equiv$ Mapping data to an infinite dimensional space and then computing an inner (dot) product

Consider any x_i



$$\exp\left(-\frac{\|x_i - x_j\|^2}{\sigma}\right)$$