SPARSE KERNEL MACHINES (CH. 7)

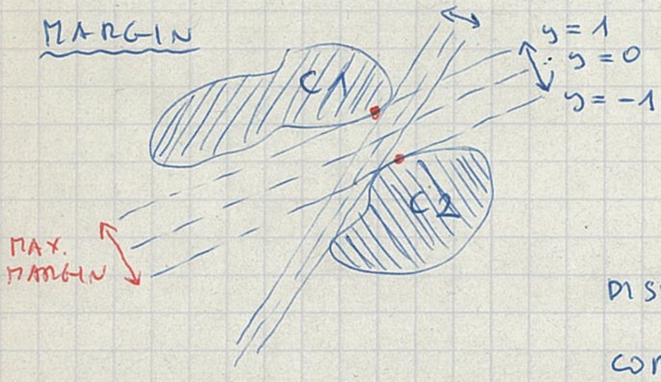
SVM – SUPPORT VECTOR MACHINE $\longrightarrow$ [CONVEX OPT. / GLOBAL OPT.

MAXIMUM MARGIN CLASSIFIER (CH. 7.1)

$K=2$ $\quad y(\underline{x}) = \underline{w}^T \phi(\underline{x}) + b \quad \{\underline{x}_n, t_n\}_{n=1}^{N} \quad t_n \begin{cases} +1 \\ -1 \end{cases}$

$$y(\underline{x}) \gtrless \phi$$

MARGIN



SMALLEST DISTANCE BETWEEN THE DECISION BOUNDARY AND ANY OF THE SAMPLEST

$\downarrow$

MAXIMIZED

DISTANCE TO HYPERPLANE $\dfrac{|y(\underline{x})|}{\|\underline{w}\|}$

CORRECT CLASSIFICATION

$$t_n y(\underline{x}_n) > \phi$$

DISTANCE OF x_n TO THE DECISION SURFACE

$$\frac{t_n y(x_n)}{\|w\|} = \frac{t_n(\underline{w}^T \phi(\underline{x}) + b)}{\|\underline{w}\|}$$

$$\arg\max_{\underline{w}, \underline{b}} \left\{ \frac{1}{\|w\|} \min_n \left[t_n(\underline{w}^T \phi(\underline{x}) + b) \right] \right\}$$

SCALING: FOR A POINT CLOSEST TO THE SURFACE $\quad t_n(\underline{w}^T \phi(\underline{x}_n) + b) = 1$

ALL POINTS $\quad t_n(\underline{w}^T \phi(\underline{x}) + b) \geq 1 \quad n = 1, \dots, N \quad (*)$

$=$ ACTIVE CONSTRAINT

$>$ INACTIVE —"—

PROBLEM: $\arg\min_{\underline{w}, b} \frac{1}{2}\|\underline{w}\|^2$ + CONSTRAINT $(*)$

QUADRATIC PROGRAMMING PROBLEM

$$L(\underline{w}, b, \underline{a}) = \frac{1}{2}\|\underline{w}\|^2 - \sum_{n=1}^{N} a_n \left\{ t_n(\underline{w}^T \phi(\underline{x}) + b) - 1 \right\}$$

LAGRANGE – MULTIPLIERS $a_n \geq \phi$