KHU-BME
Pattern Classification
Lecture 04-1

Linear Training Algorithms

2 Class Problems

Assumption: Prototypes are linearly separable
Objective: Determine linear hyperplane(s) that separate prototypes according to their classes

So find $\underline{w}^{(k)}$ and $w_{N+1}^{(k)}$

such that

$$g(y_m^{(1)}) = \underline{w}^T y_m^{(1)} > 0 \qquad \forall \underline{y}_m^{(1)} \in S_1$$

$$g(y_m^{(2)}) = \underline{w}^T y_m^{(2)} < 0 \qquad \forall \underline{y}_m^{(2)} \in S_2$$

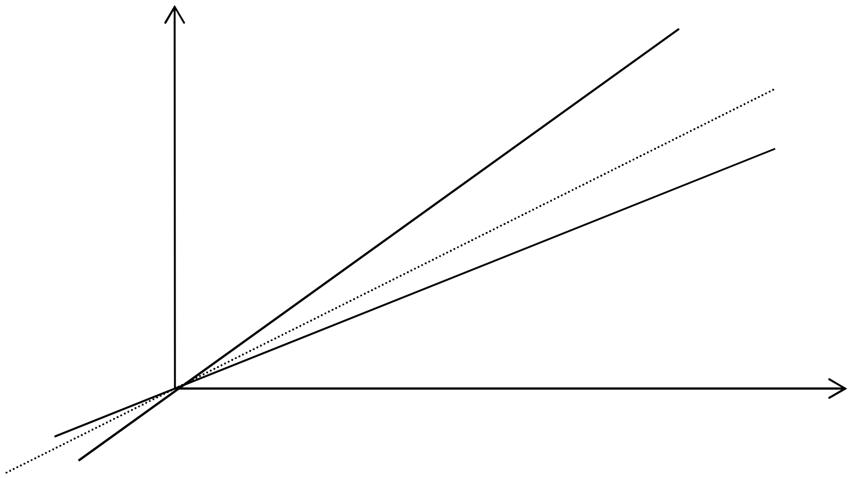
Then for unknowns, $\underline{x}$

$g(\underline{x}) = \underline{w}^T \underline{x} > 0$ -> $\underline{x} \in S_1$

$g(\underline{x}) = \underline{w}^T \underline{x} < 0$ -> $\underline{x} \in S_2$

Previously, $\underline{w}$ is given (or fixed) and $\underline{x}$ are variables

But for training, $\underline{w}$ are variables, given the prototypes $y_m^{(k)}$.



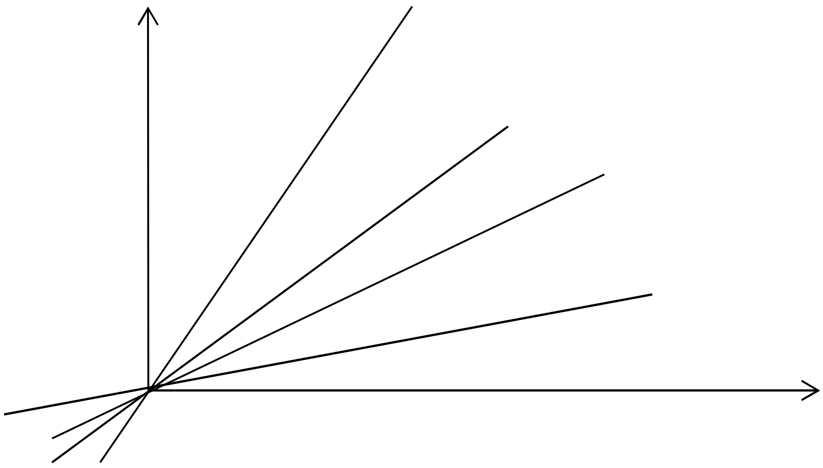
$\underline{w}^T \underline{x} = 0$ decision boundary.

Line in (x_1,x_2) space has $\underline{w}$ as parameter
Line in (w_1,w_2) space has $\underline{x}$ as parameter

$\underline{w}^T \underline{y} = 0$: Line in (w_1,w_2) space ($\underline{y}$ as parameter)

A hyperplane in the weight space $\underline{w}^T\underline{x}=0$ or $\underline{w}^T\underline{y}=0$ separates the $\underline{w}$-space into 2 regions.

Points $\underline{w}$ on one side => $\underline{w}^T\underline{y}>0$
Points $\underline{w}$ on one side => $\underline{w}^T\underline{y}<0$



Choose the weight vectors (points)

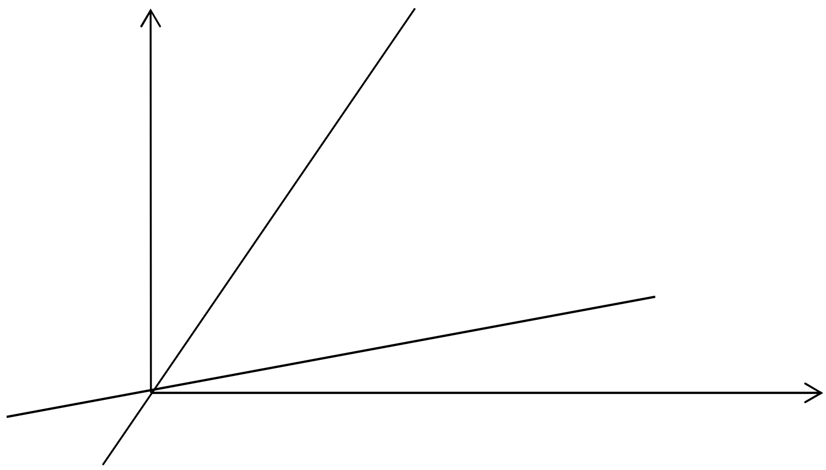
On the positive side of $\underline{y}_m^{(1)}$ hyperplanes

On the negative side of $\underline{y}_m^{(2)}$ hyperplanes

$$\underline{w}^T \underline{y}_m^{(1)} > 0 \quad \forall \underline{y}_m^{(1)}$$

$$\underline{w}^T \underline{y}_m^{(2)} < 0 \quad \forall \underline{y}_m^{(2)}$$

Safety Margin



For training,

$\underline{w}^T \underline{x} > d$ Class S_1

$\underline{w}^T \underline{x} < -d$ Class S_2

This defines the solution region with the safety margin.

Reflected Prototypes

For a 2-class problem

$\underline{w}^T \underline{y}_m^{(1)} > 0 \quad \forall \underline{y}_m^{(1)} \in S_1$

$\underline{w}^T \underline{y}_m^{(2)} < 0 \quad \forall \underline{y}_m^{(2)} \in S_2$

or $\underline{w}^T (-\underline{y}_m^{(2)}) > 0 \quad \forall \underline{y}_m^{(2)} \in S_2$

Thus we can replace all $\underline{y}_m^{(2)}$ with - $\underline{\tilde{y}}_m^{(2)}$ => reflected prototypes.

(See DHS Fig. 5.8)

$\underline{w}^T \underline{\tilde{y}}_m > 0 \quad \forall \underline{\tilde{y}}_m$

If $\underline{w}^T \underline{\tilde{y}}_m < 0$ => misclassified

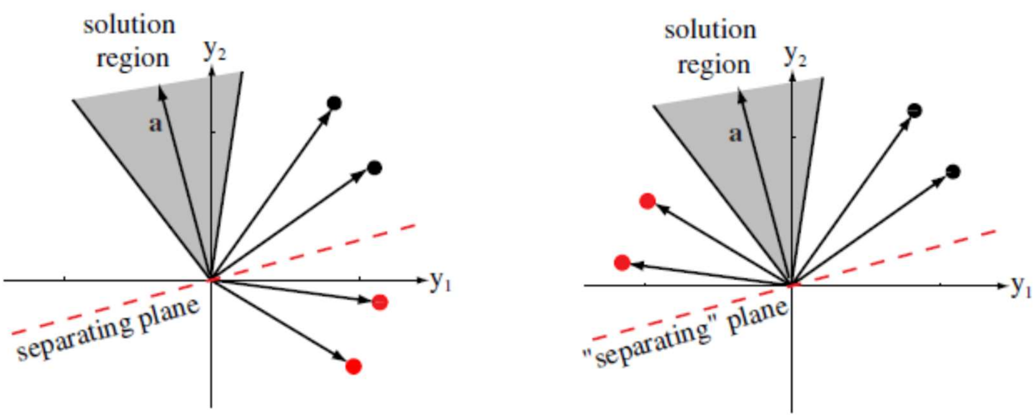


Figure 5.8: Four training samples (black for ω_1, red for ω_2) and the solution region in feature space. The figure on the left shows the raw data; the solution vectors leads to a plane that separates the patterns from the two categories. In the figure on the right, the red points have been "normalized" — i.e., changed in sign. Now the solution vector leads to a plane that places all "normalized" points on the same side.