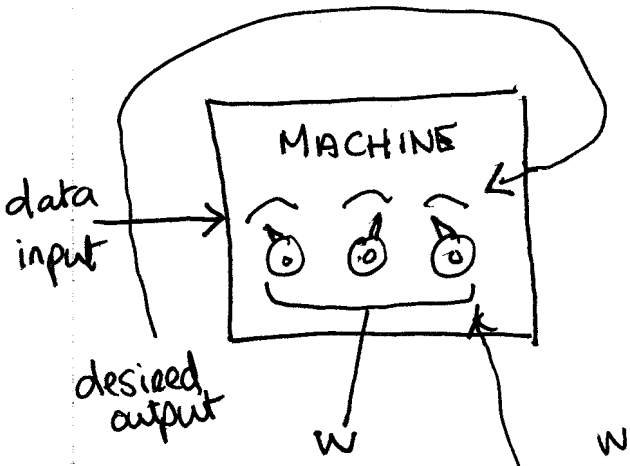
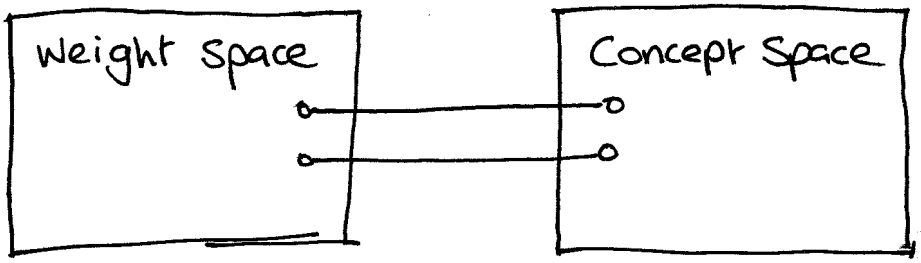
FROM LECTURE ON 4th OCTOBER

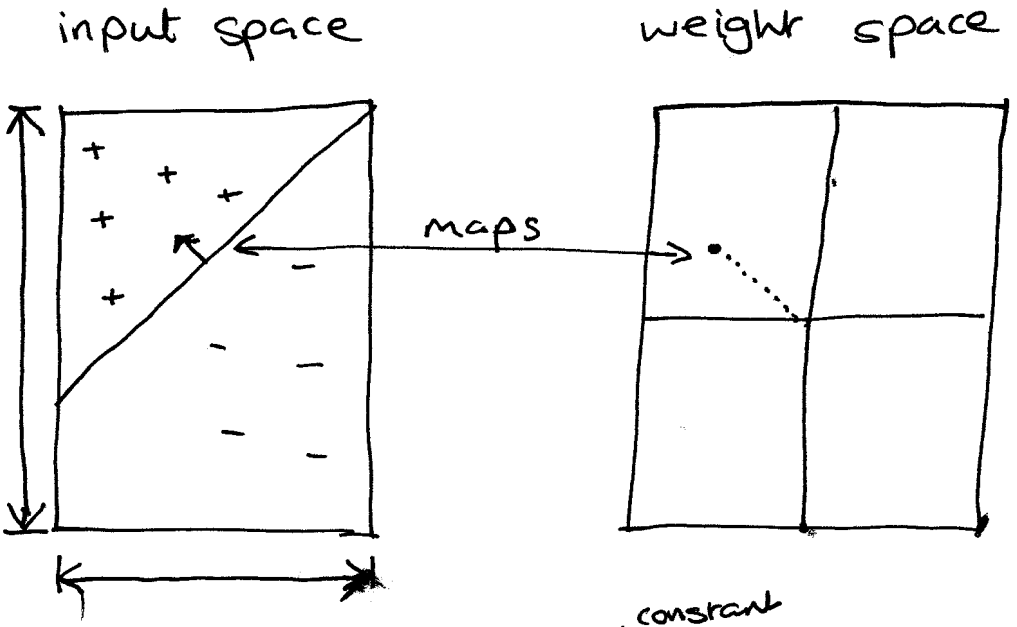


$$\left[\sum_i w_i x_i \overset{?}{>} \theta\right]$$

$$\left[\sum_i \hat{w}_i x_i \overset{?}{>} \hat{\theta}\right] \qquad \hat{\theta} = 2\theta \quad \hat{w}_i = 2w_i$$

weight space is all the possible settings (A point in weight space is one possible setting of 'knobs')



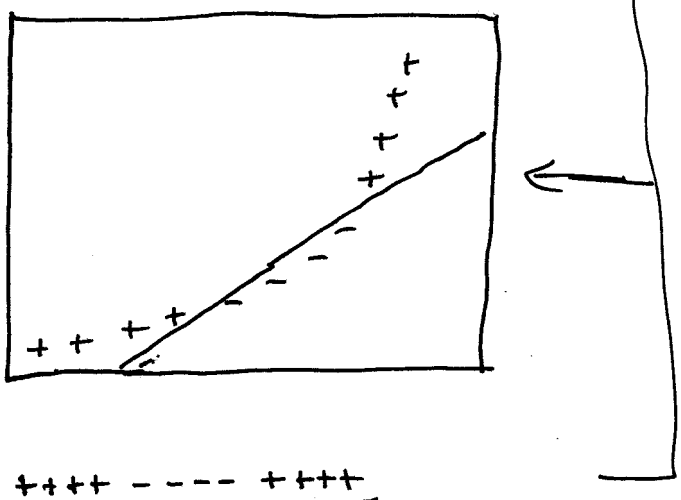


$$w_i \leftarrow w_i + (y-d)\,x_i \qquad \eta \text{ constant}$$

vector

$$\vec{w} \leftarrow \vec{w} + c\vec{x}$$

$x \longrightarrow x + x^2$)



(This method is quicker but nor as good)

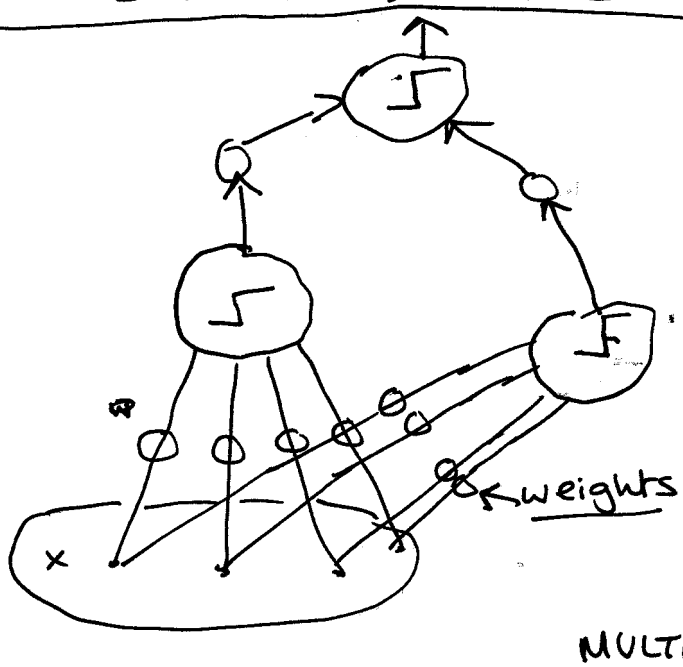
Support Vector Machines

(for pre-processing)

++++ ---- ++++

input space (mapped to higher dimension)

Do some pre-processing to map to other space so we can use linear classifier.



MULTI LAYER PERCEPTRON

(This gives best performance but takes time. Fiddly to get working.)

- Credit assignment hard

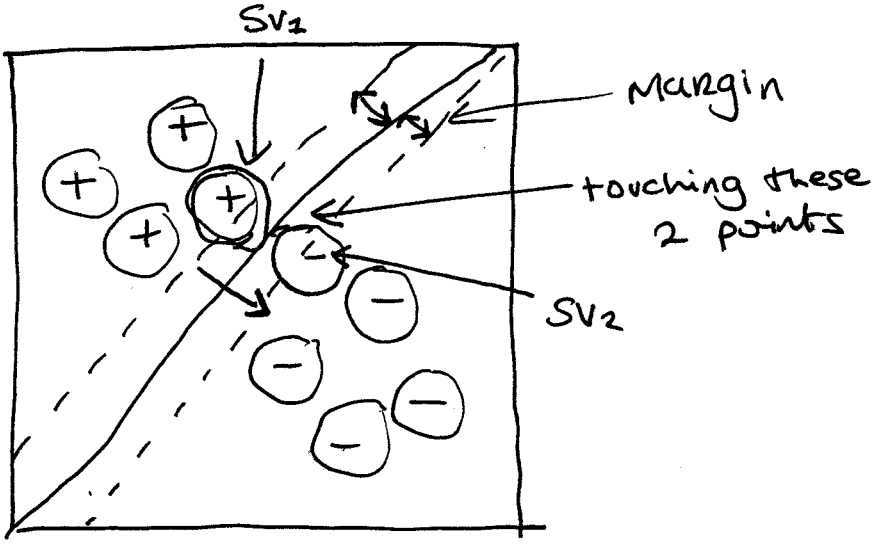
back propagation

SUPPORT VECTOR MACHINES

(3 different methods)

LINEAR SUPPORT VECTOR MACHINE

input space LTU (Linear Threshold Unit)



Many possible L.T.'s

- choose one that
  - classifies data correctly
  - it is as far away from data as possible

MAXIMUM MARGIN HYPERPLANE

- Support Vectors (SV_x) touch on the MAX MARGIN HYPERPLANE

$$x \xrightarrow{\Phi} \Phi(x)$$

$$\Phi: \mathbb{R}^5 \rightarrow \mathbb{R}^{10,000}$$

($\mathbb{R}^5$: data space; $\mathbb{R}^{10,000}$: feature space)

Computational Issues Involved:

- high dimensional feature space, computational burden
- outliers, mislabeled data.
- find MMH, represents etc.