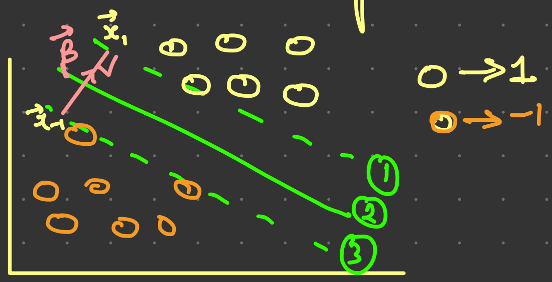
Derivation of Quadratic Constrained Quadratic form of SVM without soft constraints:

If we have the following:



All samples can be defined generally by $\boxed{(\beta\vec{x}+\beta_0)y \geq 1}$

Decision boundary is defined ②: $\beta_0 + \vec{X}\beta = 0$

①: $\beta_0 + \vec{X}\beta = 1$ → Boundary of $y=+1$

③: $\beta_0 + \vec{X}\beta = -1$ → Boundary of $y=-1$

① The points which are along the lines $\beta_0 + \vec{X}\beta = 1$ for $y=+1$
$\beta_0 + \vec{X}\beta = -1$ for $y=-1$

touching the support vectors and thereby influence it

② So any points outside that region i.e. for $y=+1$, $\beta_0 + \vec{X}\beta > 1$
for $y=-1$, $\beta_0 + \vec{X}\beta < -1$

don't influence the support vectors. i.e. non-support vectors

Now as both the support vectors are $\parallel^{lel}$ to each other, draw a $\perp^{lar}$ b/w them $\vec{\beta}$.

So, as $\beta\vec{x}_1 + \beta_0 = 1$ for a point on support vector of $y=+1$

$\beta\vec{x}_{-1} + \beta_0 = -1$ for a point on support vector of $y=-1$

Taking the difference:

$$\beta(\vec{x}_1 - \vec{x}_{-1}) = 2$$

$$\frac{2}{\beta} = (\vec{x}_1 - \vec{x}_{-1}$$

Substitution

Also from the figure

$$\frac{\vec{\beta}}{\|\beta\|}\left[\vec{x}_1 - \vec{x}_{-1}\right] = \text{Width}(W)$$

$$\boxed{\text{Width}(W) = \frac{2}{\|\beta\|}}$$

To maximise the margin, width must be maximised $\Rightarrow$ Minimise $\|\beta\|$.

So taking average absolute squation:

Must include condition for all samples:

Objective Function: $\underset{\beta}{\operatorname{argmin}} \frac{1}{2} \|\beta\|^2$

Constraint: $(\beta_0 + \beta\vec{x})y \geq 1$ (for all training samples)