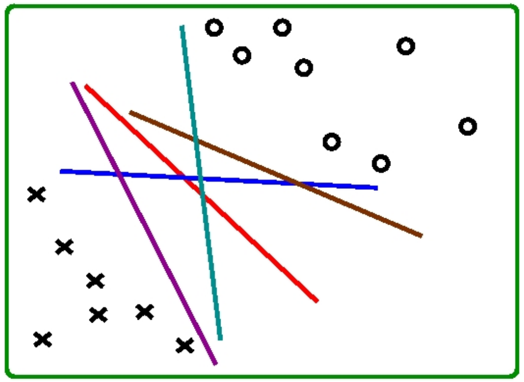
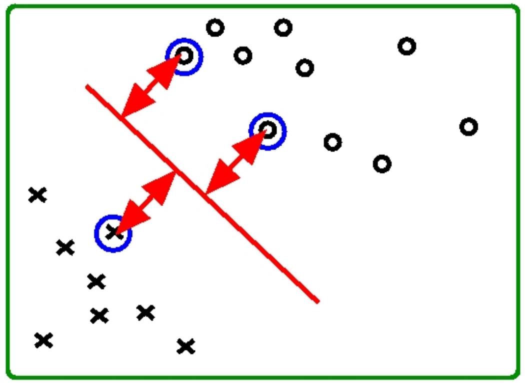
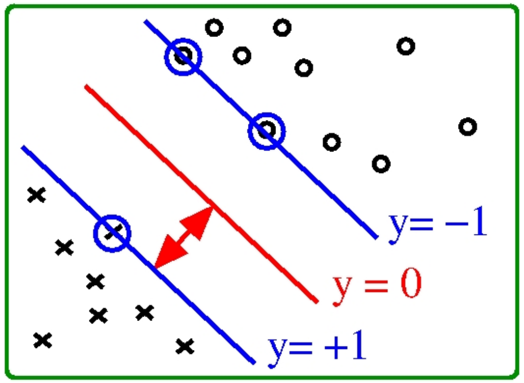


- Infinite possibilities
- Choose acc to some optimisation criterion



- min dist of a point from the decision boundary
- $\triangleq$ margin, to maximise

- Implicit form of the eqn of a line $w_2 x_2 + w_1 x_1 + b = 0$
- The two intercept form $\frac{x_2}{(-b/w_2)} + \frac{x_1}{(-b/w_1)} = 1$
- The slope-intercept form $x_2 = (\frac{-w_1}{w_2}) x_1 + (\frac{-b}{w_2})$
- Take-home point#1: $\mathbf{w}$ determines the slope
- Take-home point#2: b: scaled distance from the origin. Why? $\frac{b}{||\mathbf{w}||}$ is the distance from the origin.



- Margin: min dist b/w decision boundary & any sample
- symmetric, by defn above
- Don't worry about $y = \pm 1$, *yet*
- Aim: Maximise this margin
- The location of the boundary: determined by a small subset of the data points: Support Vectors
- Decision surface: $y(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}) + b = 0$
- Perp dist of $\mathbf{x}$ from $y(\mathbf{x}) = 0$: is given by $\frac{|y(\mathbf{x})|}{||\mathbf{w}||}$
- We want correct classification i.e., $t_i\, y(\mathbf{x}_i) > 0\ \forall i$
- $|y(\mathbf{x}_i)| = +y(\mathbf{x}_i)$ when $y(\mathbf{x}_i) > 0\ (t_i = +1)$
- $|y(\mathbf{x}_i)| = -y(\mathbf{x}_i)$ when $y(\mathbf{x}_i) < 0\ (t_i = -1)$
- $\Longrightarrow\ |y(\mathbf{x})| = t_i\, y(\mathbf{x}_i)$, perp dist = $\frac{t_i\, y(\mathbf{x}_i)}{||\mathbf{w}||}$
- Max margin: $\boxed{\arg\max_{\mathbf{w},b}\{\frac{1}{||\mathbf{w}||}\min_i\{t_i\,[\mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i)+b]\}\}}$

- margin: $\triangleq$ min dist frm a point in either class: sym

Interpretation of: $\arg\max_{\mathbf{w},b}\{\frac{1}{||\mathbf{w}||}\min_i\{t_i\,[\mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i)+b]\}\}$

$$= \underbrace{\arg\max_{\mathbf{w},b}\{\min_i \overbrace{\underbrace{\frac{\{t_i\ \overbrace{[\mathbf{w}^T\overbrace{\boldsymbol{\phi}(\mathbf{x}_i)}^{or,\ \mathbf{x}_i}+b]}^{\text{eqn of line}_{y(\mathbf{x}_i)=0}}\}}{||\mathbf{w}||}}_{\text{dist of } \mathbf{x}_i \text{ from } y(\mathbf{x}_i)=0}}^{} \}}_{\triangleq\ margin}$$

to maximise this margin

- Margin = $\pm\frac{1}{||\mathbf{w}||}$: particularly elegant
- This is just a scaling. Scaling $\mathbf{w}$ and b by κ leave the margin unchanged (shown later)

Home Page

Title Page

Contents

Page 7 of 41

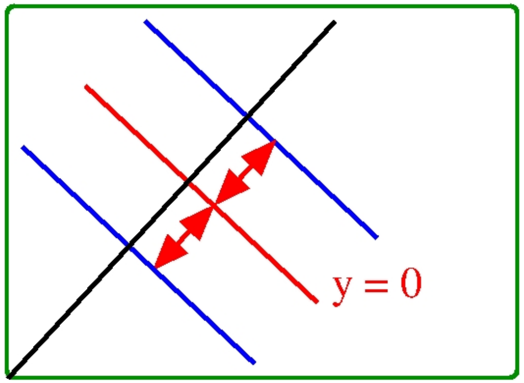
Go Back

Full Screen

Close

Quit

Three Men in a Boat: Three Lines, Eqns



- $y = 0$: decision boundary
- Implicit form
  $y(\mathbf{x}_i) = \mathbf{w}^T\mathbf{x}_i + b = 0$
- Slope: $\mathbf{w}$, dist from origin $\frac{b}{||\mathbf{w}||}$
- $||$ lines: same slope, diff dist
- 2 lines normalised dist $\frac{1}{||\mathbf{w}||}$
- 'Near line': same slope, closer than dec boundary
- closer by $\frac{1}{||\mathbf{w}||}$: from origin: $\frac{b}{||\mathbf{w}||} - \frac{1}{||\mathbf{w}||}$:coeff = $(b-1)$
- $\mathbf{w}^T\mathbf{x}_i + (b-1) = 0 \implies \mathbf{w}^T\mathbf{x}_i + b = +1 \implies y = +1$
- 'Far line': same slope, farther than dec boundary
- farther by $\frac{1}{||\mathbf{w}||}$: from origin: $\frac{b}{||\mathbf{w}||} + \frac{1}{||\mathbf{w}||}$:coeff = $(b+1)$
- $\mathbf{w}^T\mathbf{x}_i + (b+1) = 0 \implies \mathbf{w}^T\mathbf{x}_i + b = -1 \implies y = -1$

Home Page

Title Page

Contents

◀◀ ▶▶

◀ ▶

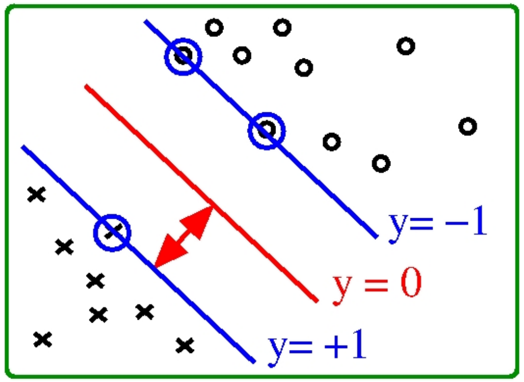
Page 8 of 41

Go Back

Full Screen

Close

Quit

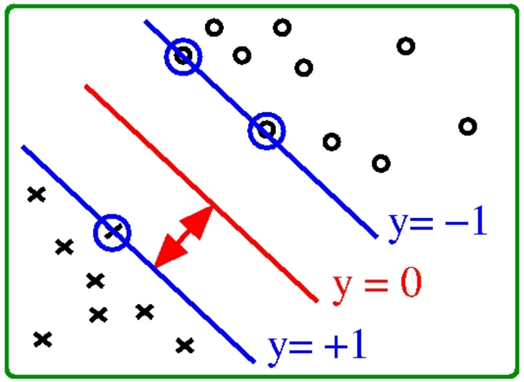


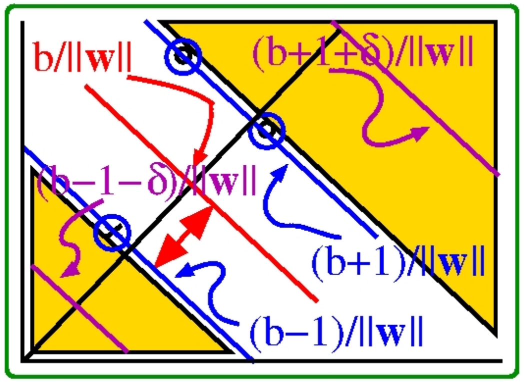
- Maximum margin solution:

$\arg\max_{\mathbf{w},b}\{\frac{1}{||\mathbf{w}||}\min_i\{t_i\ [\mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i)+b]\}\}$

- 'min' comes from 'margin'
- Find $\mathbf{w}, b$ to max the margin
- Now let us look at $y = \pm 1$
- Consider $\boldsymbol{\phi}(\mathbf{x}_i) = \mathbf{x}_i$ (simplicity, no feature xform)
- $y(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b$ is a hyperplane/line in $\mathbf{x}-$ space
- $\mathbf{w}$ measures the slope/inclination. Why?
- $w_2x_2 + w_1x_1 + b = 0 : \frac{x_2}{(-b/w_2)} + \frac{x_1}{(-b/w_1)} = 1$
- $\frac{b}{||\mathbf{w}||}$: distance from the origin. Vary b: $||$ lines
- If somehow we know the direction $\mathbf{w}$, fit a red line equidistant from the two lines: decision boundary
- How do we know? Oracle/QP solver for $\mathbf{w}$ (& b)
- The distance of a point from the decision boundary is unchanged on a scaling of $\mathbf{w}$ & b by κ each

- The distance of a point from the decision boundary is unchanged on a scaling of $\mathbf{w}$ & b by κ each
- $= \frac{t_i\ \kappa\mathbf{w}^T\mathbf{x}_i+\kappa b}{\kappa||\mathbf{w}||} = \frac{t_i\ \mathbf{w}^T\mathbf{x}_i+b}{||\mathbf{w}||}$ (property)
- 'Nice' formulation: Consider total margin = $2/||\mathbf{w}||$





- Maximum margin solution:
  $\arg\max_{\mathbf{w},b}\{\frac{1}{||\mathbf{w}||}\min_i\{t_i\ \mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i)+b\}\}$
- 'min' comes from 'margin'
- Find $\mathbf{w}, b$ to max the margin
- Now, the $y = \pm 1$ part:
- 2 blue lines @dist$\pm\frac{1}{||\mathbf{w}||} \implies$
- coeff ± 1: $\mathbf{w}^T\mathbf{x} + (b \pm 1) = 0$
- $y = \mathbf{w}^T\mathbf{x} + b = \mp 1$. b +ve
- 'near' line: $y = \mathbf{w}^T\mathbf{x} + b = +1$
- 'far' line: $y = \mathbf{w}^T\mathbf{x} + b = -1$