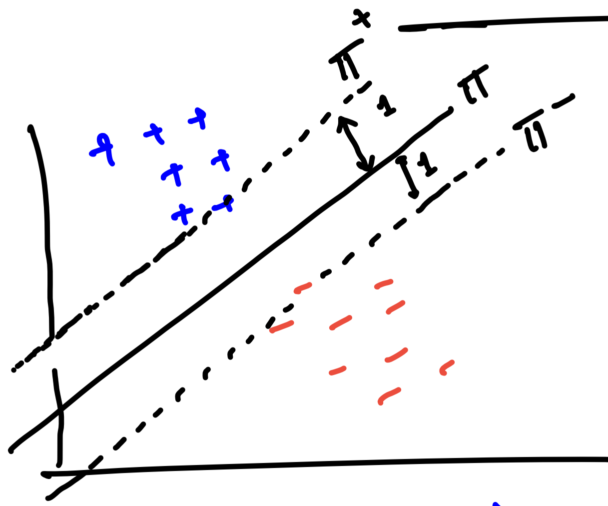
SVM - Soft Margin

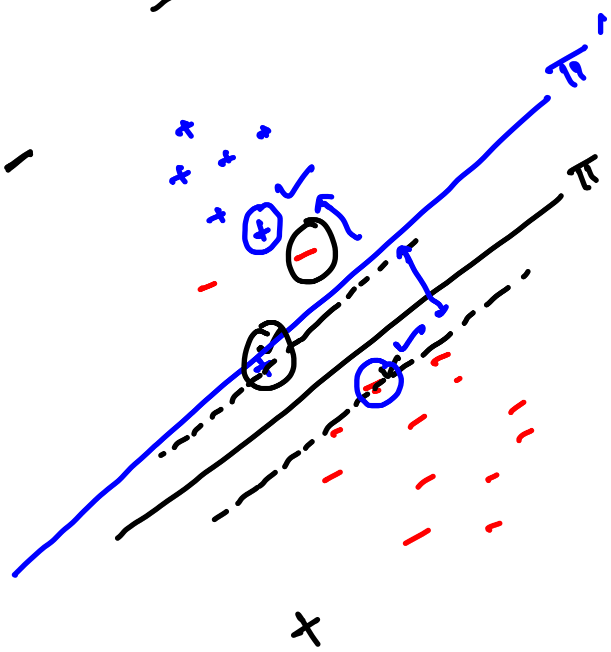
30/09/24

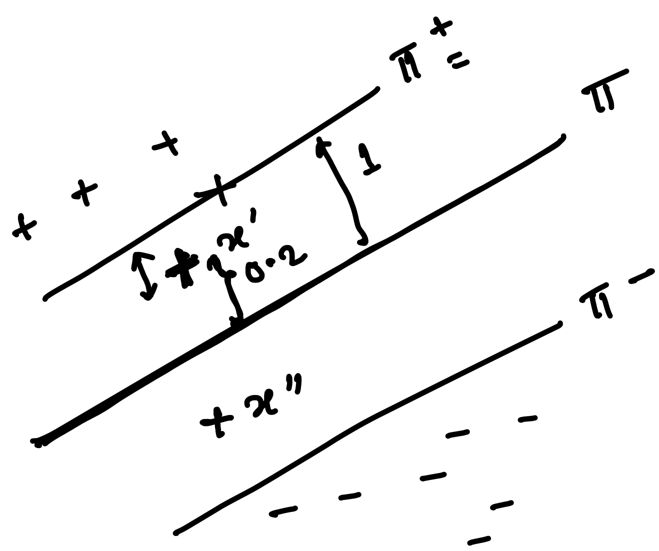


$$\min_{w,b} \frac{1}{2}\|w\|^2$$

s.t. $y^i(w^T x_i + b) \geq 1 \quad \forall i$

$$\min_{w,b} \frac{1}{2}\|w\|^2 + \underline{C}\,(\text{loss } fn^n)$$





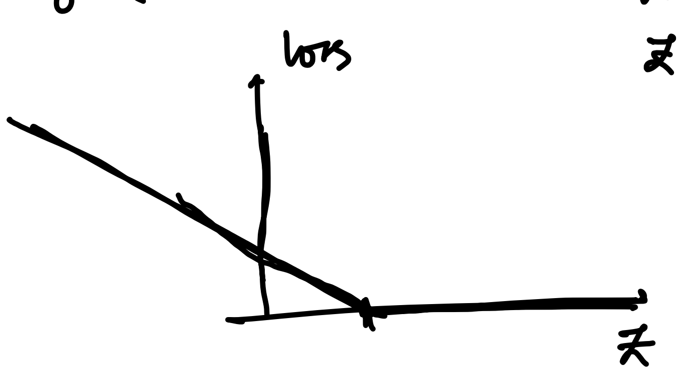
$y'(w^T x' + b) = 0.2$

$y''\underbrace{(w^T x'' + b)} = -0.2 \qquad 1 - (-0.2) = 1.2$

$y^i(w^T x_i + b) = Z_i$

$Z_i \geq 1$ loss 0

$Z_i < 1$ loss $1 - Z_i$



Hinge loss. $= \xi_i = \max(0, 1 - Z_i)$

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i \leftarrow \text{slack variable.}$$

s.t. $y_i(w^T x_i + b) \geq 1 - \xi_i$

$\xi_i \geq 0$

Logistic Reg

min log loss + $\lambda\|w\|^2$

C : hyperparameter

say $C = 1000$

$\frac{1}{2}\|w\|^2 = 200$

$\sum \xi_i = 2$ – small loss

$200 + 2 \times 1000$

$= 200 + 2000$

$= \underline{2200}$ $\Rightarrow$ overfit

say: $C = 0.001$

$200 + \underline{0.001 \times 200}$

$\uparrow$ margin

underfit

Bias and variance

Regression Setup.

$y_i \in \mathbb{R}$

$\underline{D} : \{(x_1, y_1), (x_2, y_2) \ldots, (x_n, y_n)\}$

$P(x,y) = P(y|x)\,P(x)$

Sampled i.i.d

Expected value of y: $\bar{y}(x) = E_{y|x}(y) = \int y\, P(y|x)\, dy.$

$h_D \leftarrow A(D)$

Expected Test error: given h_D

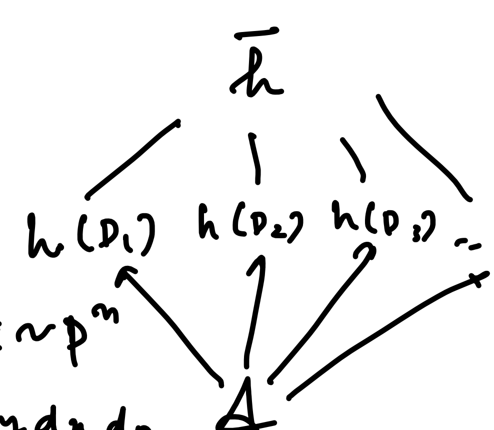
$$E_{(x,y)\sim D}\left[(h_D(x) - y)^2\right] = \int_x \int_y (h_D(x) - y)^2\, P(x,y)\, dx\, dy$$

$h_D(x)$ = R.V. as D is R.V. $\qquad D \sim p^m$

Avg hypothesis

$$\bar{h} = E_{D\sim p^m}[A(D)] = \int_D h_D\, P(D)\, dD$$

'Weak law of large Number'



Expected error of A

$$E_{\substack{(x,y)\sim D \\ D\sim p^m}} \underline{\underline{\left[(h_D(x) - y)^2\right]}} = \int_D \int_x \int_y (h_D(x) - y)^2\, P(x,y)\, P(D)\, dy\, dx\, dD$$

$$\Rightarrow E\left[\left[(h_D(x) - \bar{h}(x)) + (\bar{h}(x) - y)\right]^2\right]$$

$$\Rightarrow E_{\substack{x \\ D}}\left[\left[h_D(x) - \bar{h}(x)\right]^2\right] + E_{\substack{x \\ y}}\left[\left[\bar{h}(x) - y\right]^2\right] + 2\, E_{\substack{(x,y) \\ D}}\left[(h_D(x) - \bar{h}(x))(\bar{h}(x) - y)\right]$$

$= 0$

$$E_{xy}\left[\left[E_D\left(h_D^{(x)} - \bar{h}(x)\right]\left(\bar{h}(x) - y\right)\right]\right.$$

$$\Rightarrow E_{xy}\left[\underbrace{\left(\underset{\bar{h}(x)}{E_D\left[h_D(x)\right]} - \bar{h}(x)\right)}_{0}\left(\bar{h}(x) - y\right)\right]$$

$$\underbrace{\boxed{E_{x,D}\left[\left[h_D(x) - \bar{h}(x)\right]^2\right]}}_{A} + \underbrace{E_{(x,y)}\left[\left[\bar{h}(x) - y\right]^2\right]}_{B}$$

$$E_{(x,y)}\left[\left[\left(\bar{h}(x) - \bar{y}(x)\right) + \left(\bar{y}(x) - y\right)\right]^2\right]$$

$$= \underbrace{E\left[\left(\bar{h}(x) - \bar{y}(x)\right)^2\right]}_{B1} + \underbrace{E\left[\left[\bar{y}(x) - y\right]^2\right]}_{B2} + 2 . a . b$$

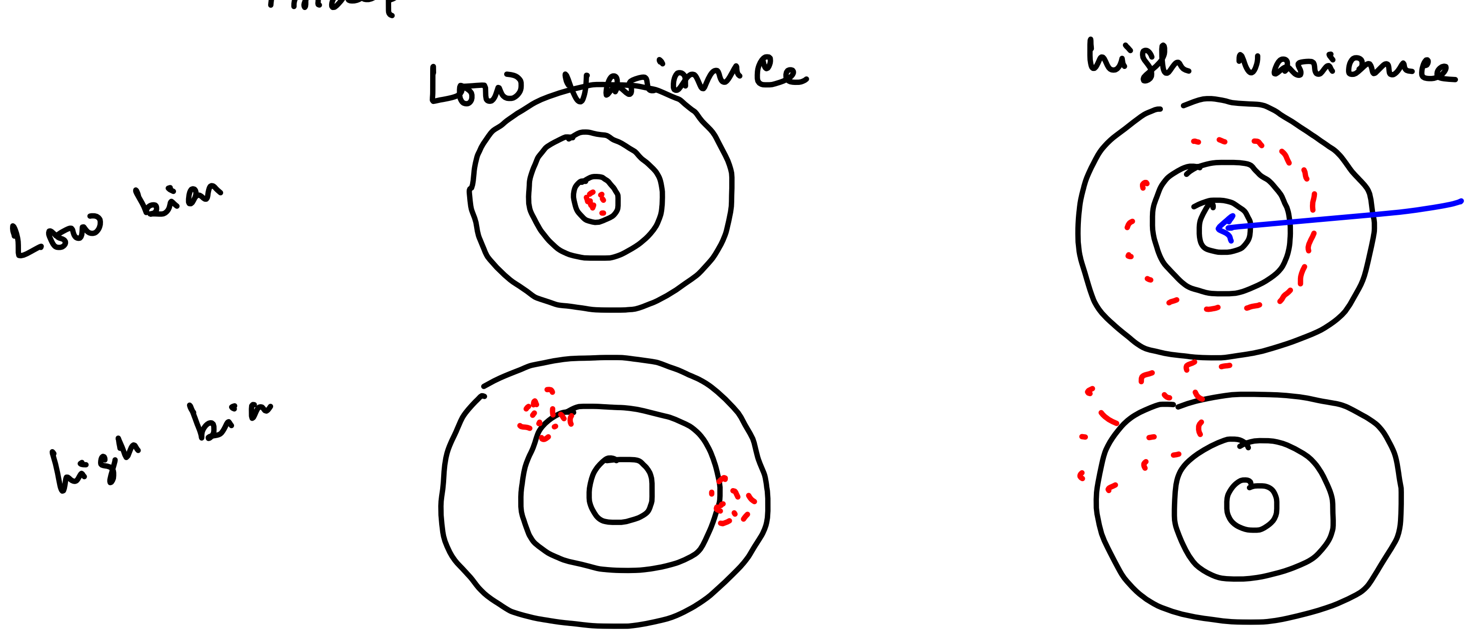
$$2 . a . b : \quad E_x\left[\left[E_{y|x}\left(\bar{y}(x) - y\right)\right]\left[\bar{h}(x) - \bar{y}(x)\right]\right]$$

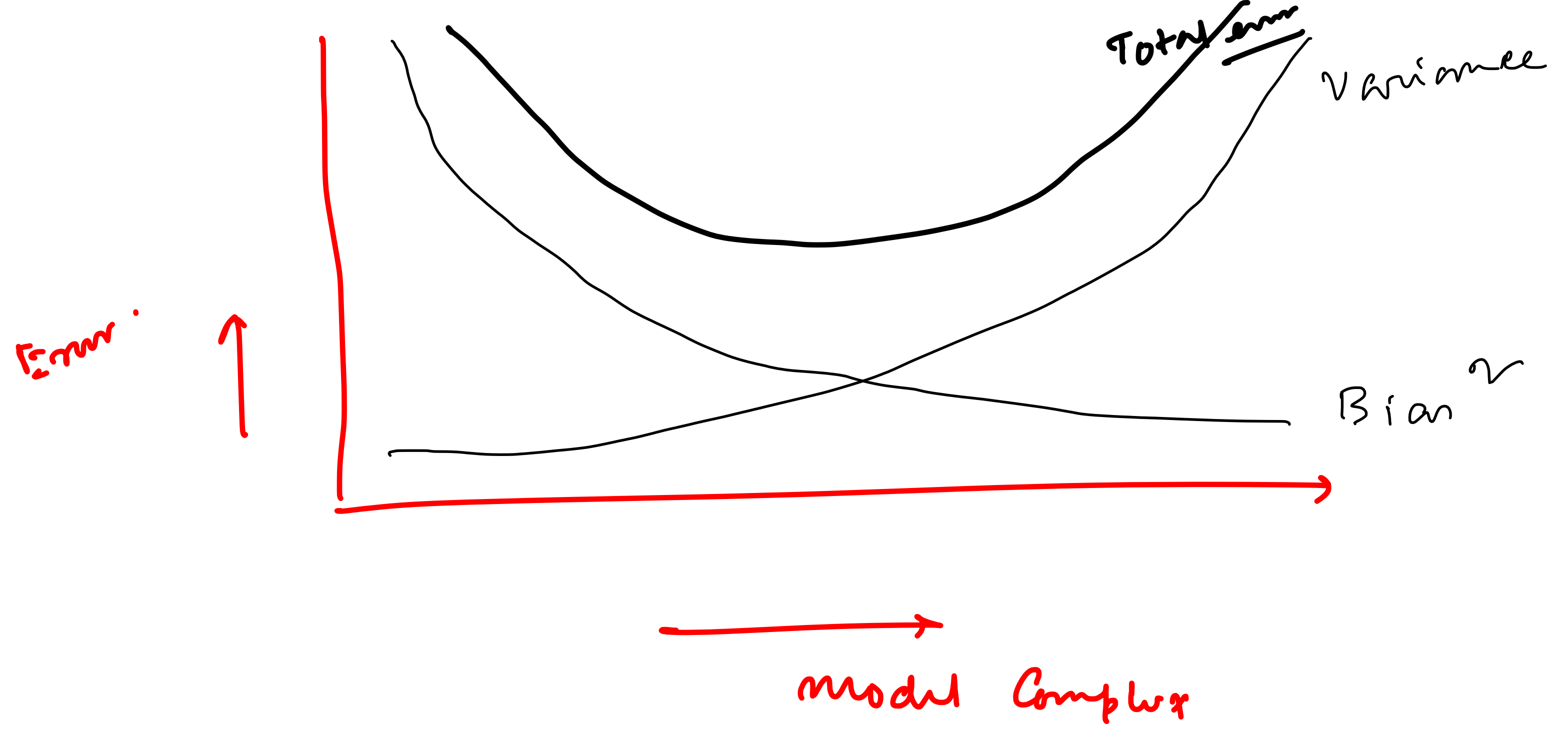
$$\Rightarrow \underbrace{E_x\left[\bar{y}(x) - \underset{\bar{y}(x)}{E_{y|x}(y)}\right]\left[\quad '' \quad\right]}_{0}$$

Expected error of A:

$$A + B1 + B2$$

$$= \underbrace{E_{x,D}\left[\left[h_D(x) - \bar{h}(x)\right]^2\right]}_{\text{Variance, independent of } y} + \underbrace{E_x\left[\left(\bar{h}(x) - \bar{y}(x)\right)^2\right]}_{\text{Bias}^2} + \underbrace{E_{x,y}\left[\left[\bar{y}(x) - y\right]^2\right]}_{\text{Noise, data-intrinsic Noise}}$$





Total error
Variance
Bias²
Error
Model Complex