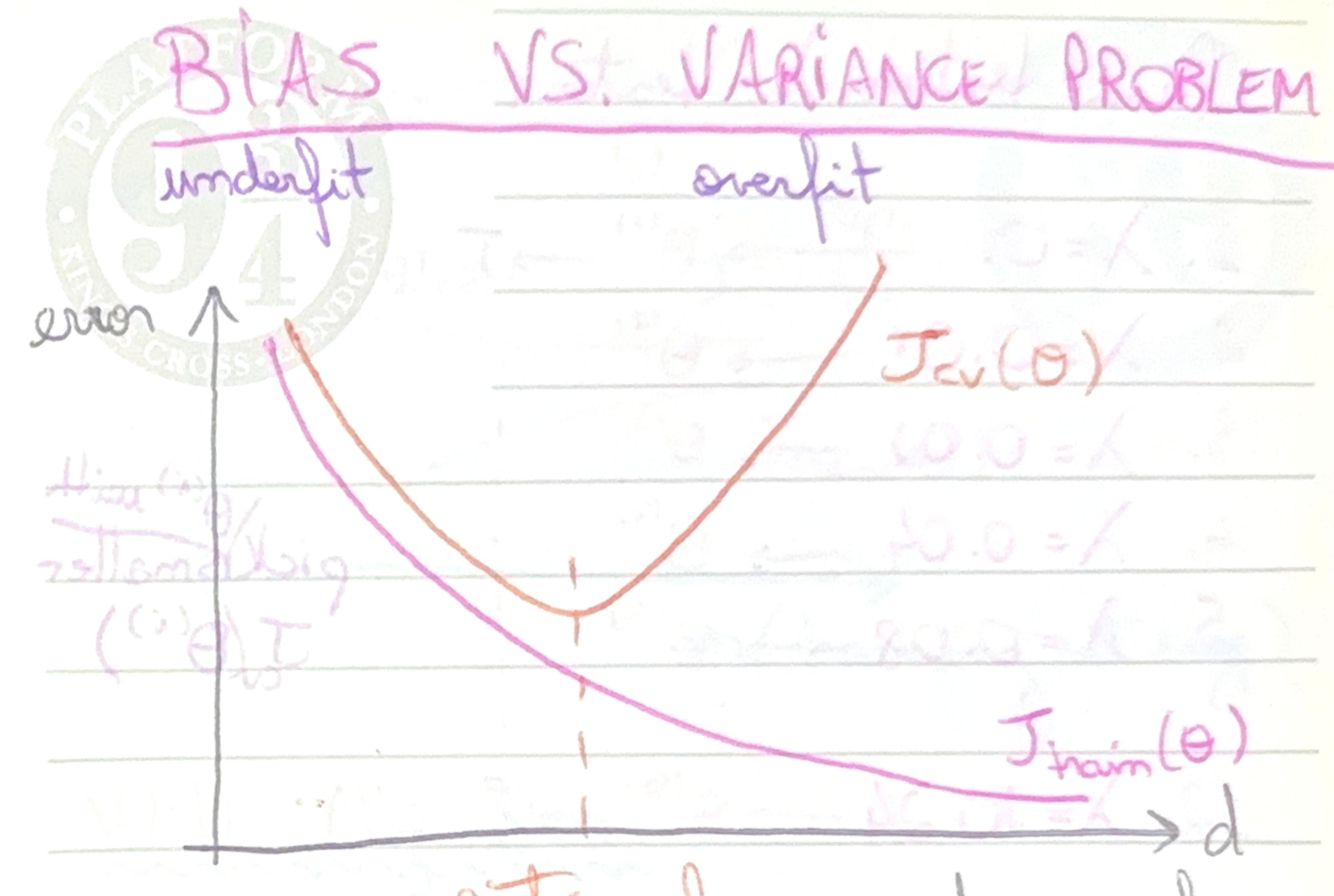


Agora, usaremos o training set para minimizar  $J(\theta)$  e obter os parâmetros  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(10)}$  para os modelos que queremos testar.

Depois, usamos esses parâmetros para calcular qual modelo tem o menor custo:  $J(\theta^{(1)}), J_{cv}(\theta^{(1)}), \dots, J_{cv}(\theta^{(10)})$ .

Com o modelo que forneciu o menor custo no cross validation set, nós vamos estimar o erro como visto na generalização por meio do cálculo do erro cometido no test set.

OBS: geralmente teremos  $J_{cv}(\theta) < J_{test}(\theta)$  porque no cross validation set nós ajustamos um feature a mais (o grau "d" do polinômio).



$(\theta)$ , degree of polynomial

problema de bias : high  $J_{train}$   
UNDERFIT  
high  $J_{cv}$  ( $J_{train} \approx J_{cv}$ )

problema de variancia: low  $J_{train}$   
OVERFIT  
high  $J_{cv}$   
( $J_{cv} \gg J_{train}$ )

small  $\lambda \rightarrow$  prone to OVERFIT  
big  $\lambda \rightarrow$  prone to UNDERFIT