

Goals of lecture:

- what is a supervised task?
- what is linear regression?
- how to find a direct solution to linear regression?

Linear regression

$(x^{(i)}, t^{(i)})$   
 ↑            ↑  
 input      label/target  
 (e.g., size) (e.g., price)

nb points in the dataset  $N$   
 Dataset  $\{(x^{(i)}, t^{(i)})\}_{i=1}^N$

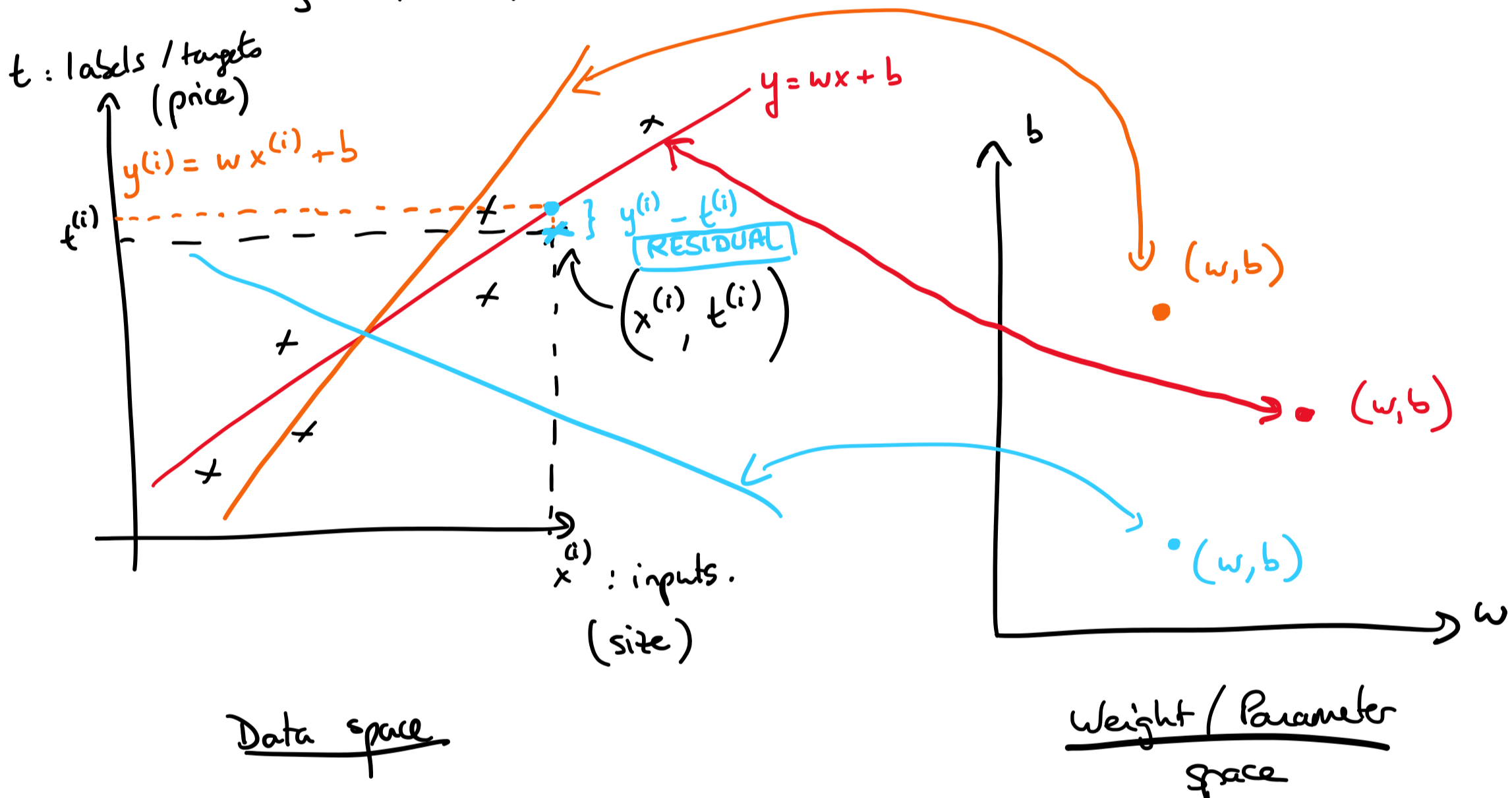
Goal of linear regression: predict  $t$  given  $x$

Model: linear model

$y = wx + b$   
 ↑            ↑            ↑  
 output of our model (prediction)    weight    input bias

output is a linear function of input

both  $w, b$  are parameters of the model.  
 settings of the parameters are called hypotheses.



Loss function

$\mathcal{L}(y, t) = \frac{1}{2} (y - t)^2$   
 ↑            ↑            ↑  
 model prediction    target for an input    residual

Goal learning: make residual as small as possible.

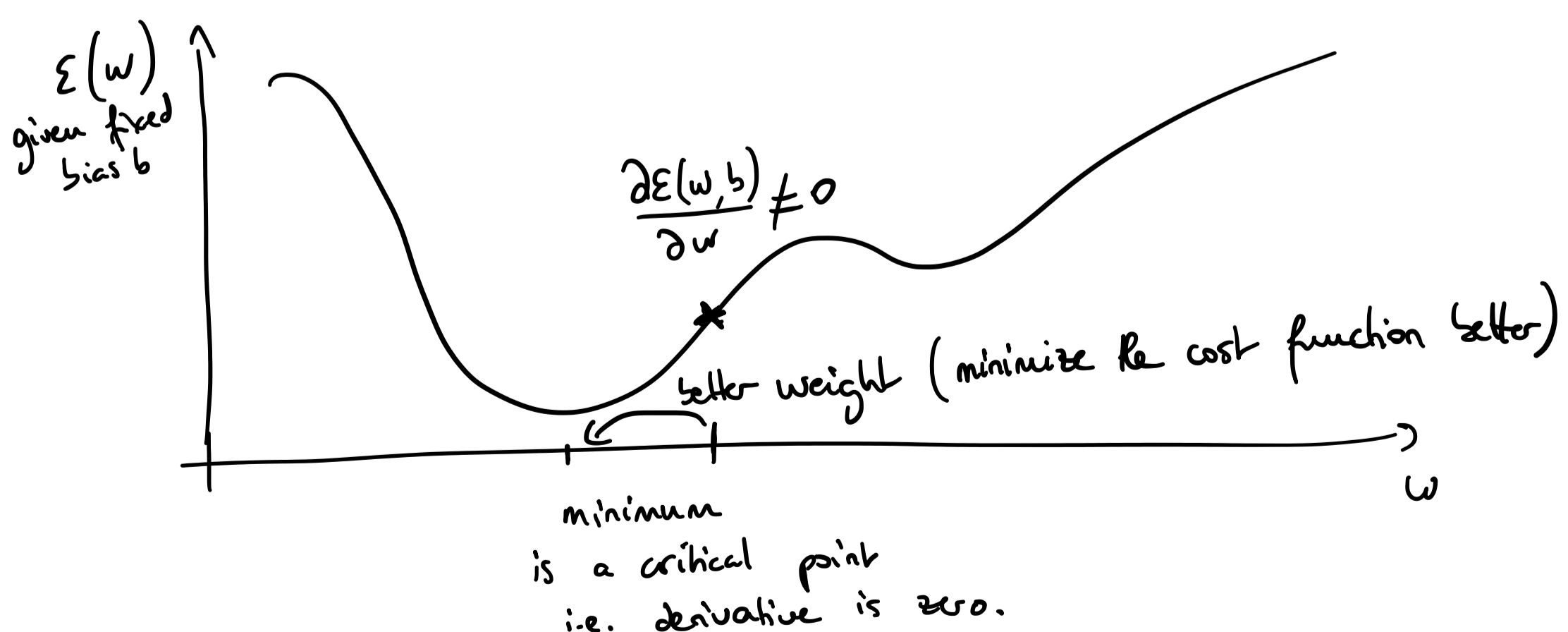
Define the loss of the model over a training set

$\mathcal{E}(w, b) = \frac{1}{2N} \sum_{i=1}^N (y^{(i)} - t^{(i)})^2$   
 ↑            ↑            ↑  
 parameters of the model    nb of training examples    residual

$= \frac{1}{2N} \sum_{i=1}^N (wx^{(i)} + b - t^{(i)})^2$

Finding the model parameters: training

Goal: minimize the loss function  $\mathcal{E}(w, b)$



Multivariate cost function (depends on both  $w, b$ )

$\frac{\partial \mathcal{E}}{\partial w} = 0$

partial derivative wrt. weight

$\frac{\partial \mathcal{E}}{\partial b} = 0$

partial derivative wrt. bias