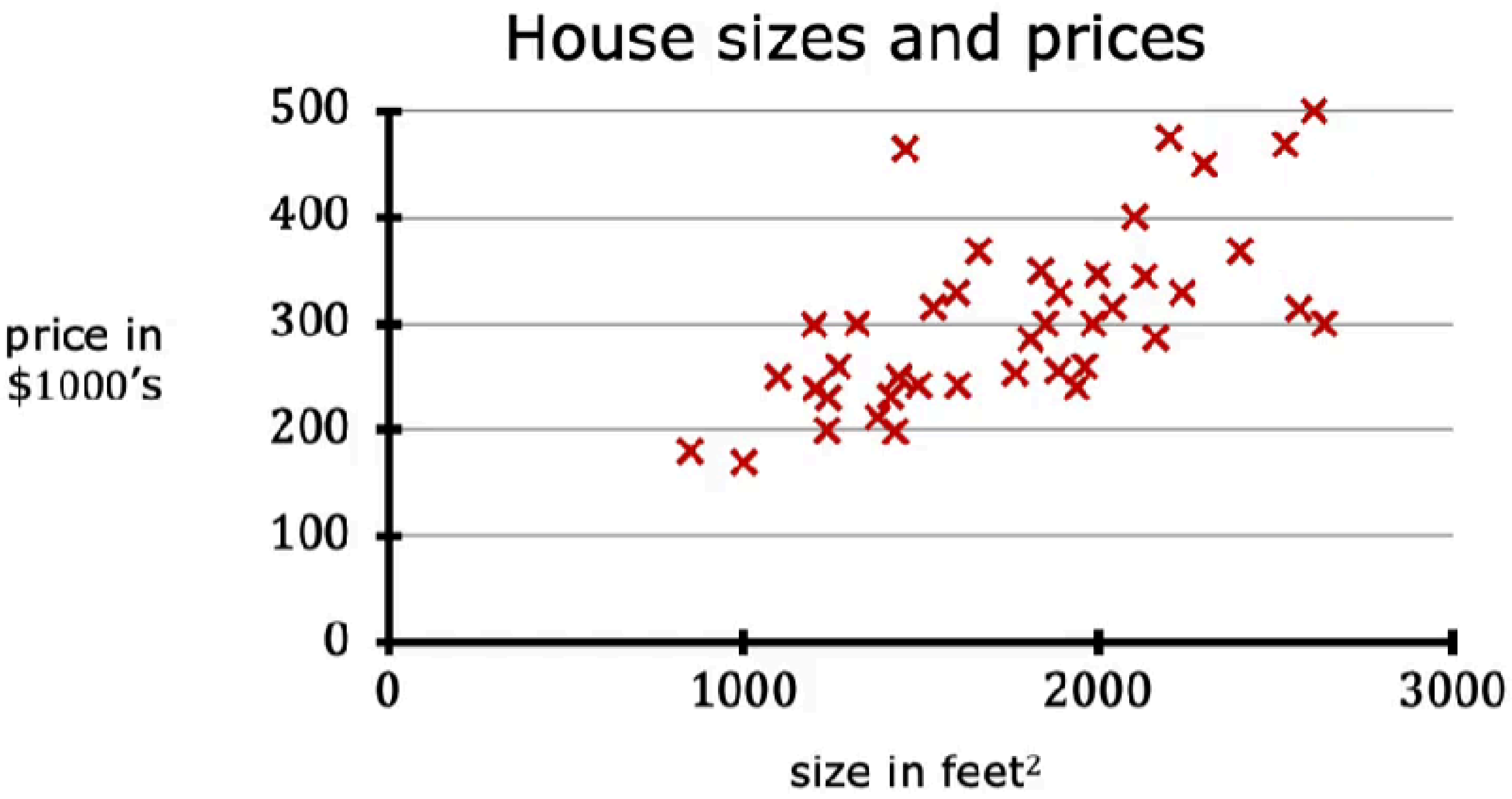
Linear Regression

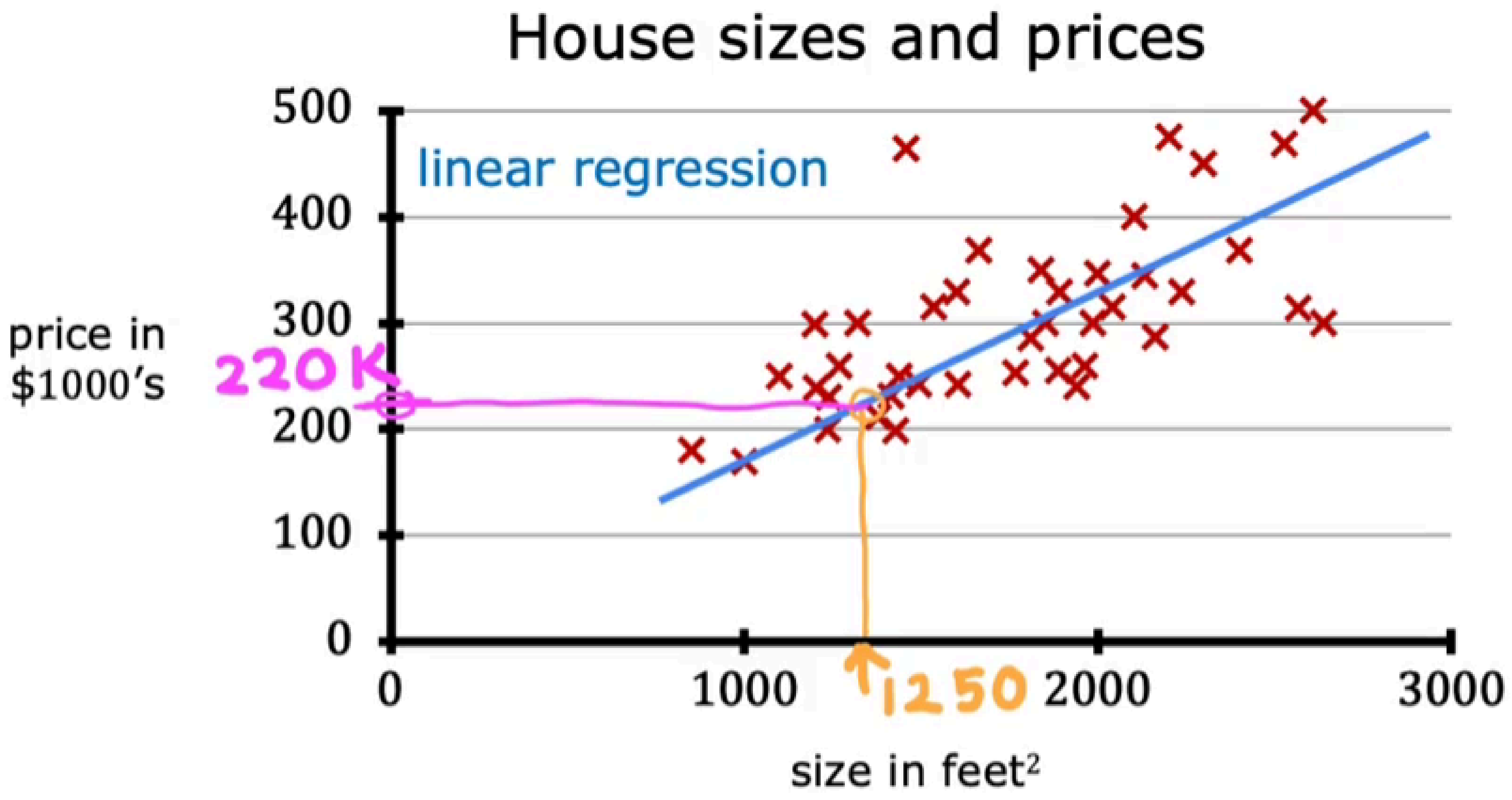
What is Linear Regression?



What is Linear Regression?

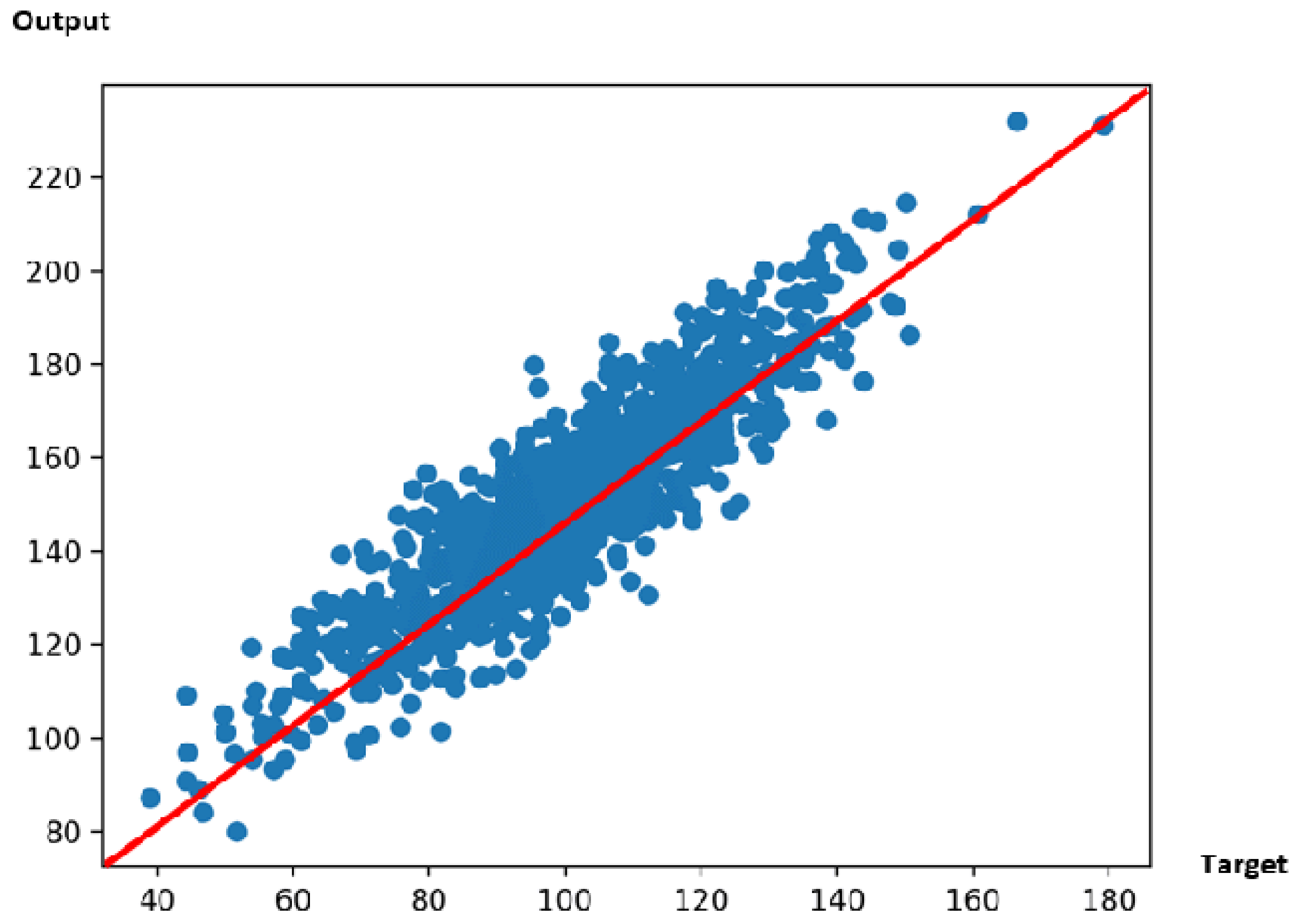
What is Linear Regression?

What is Linear Regression?



$$f_{w,b}(x) = wx + b$$

What is Linear Regression?



Notations

size in feet^2	price in \$1000's
2104	400
1416	232
1534	315
852	178
...	...
3210	870

Notation:

x = "input" variable
feature

y = "output" variable
"target" variable

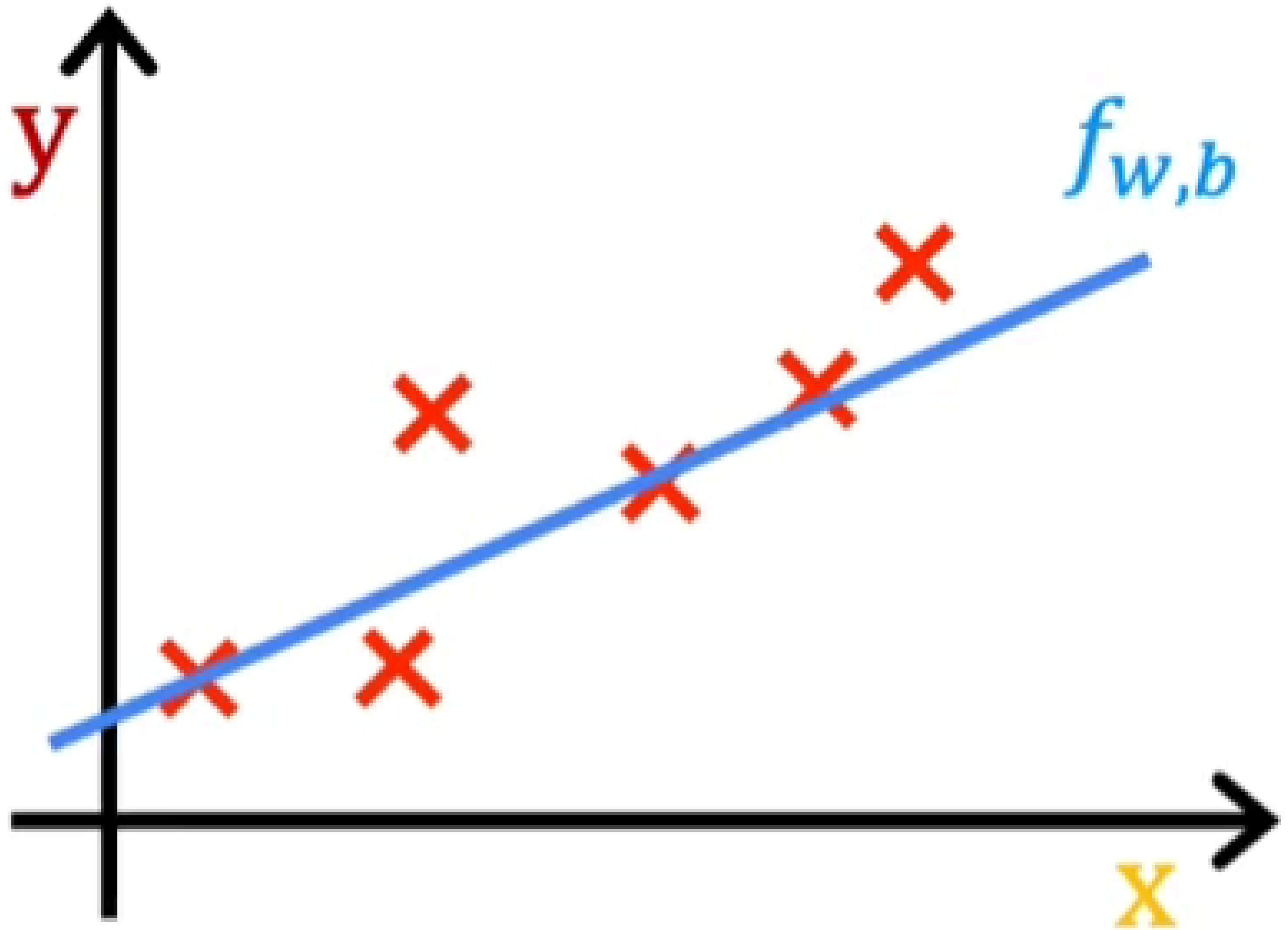
m = number of training examples

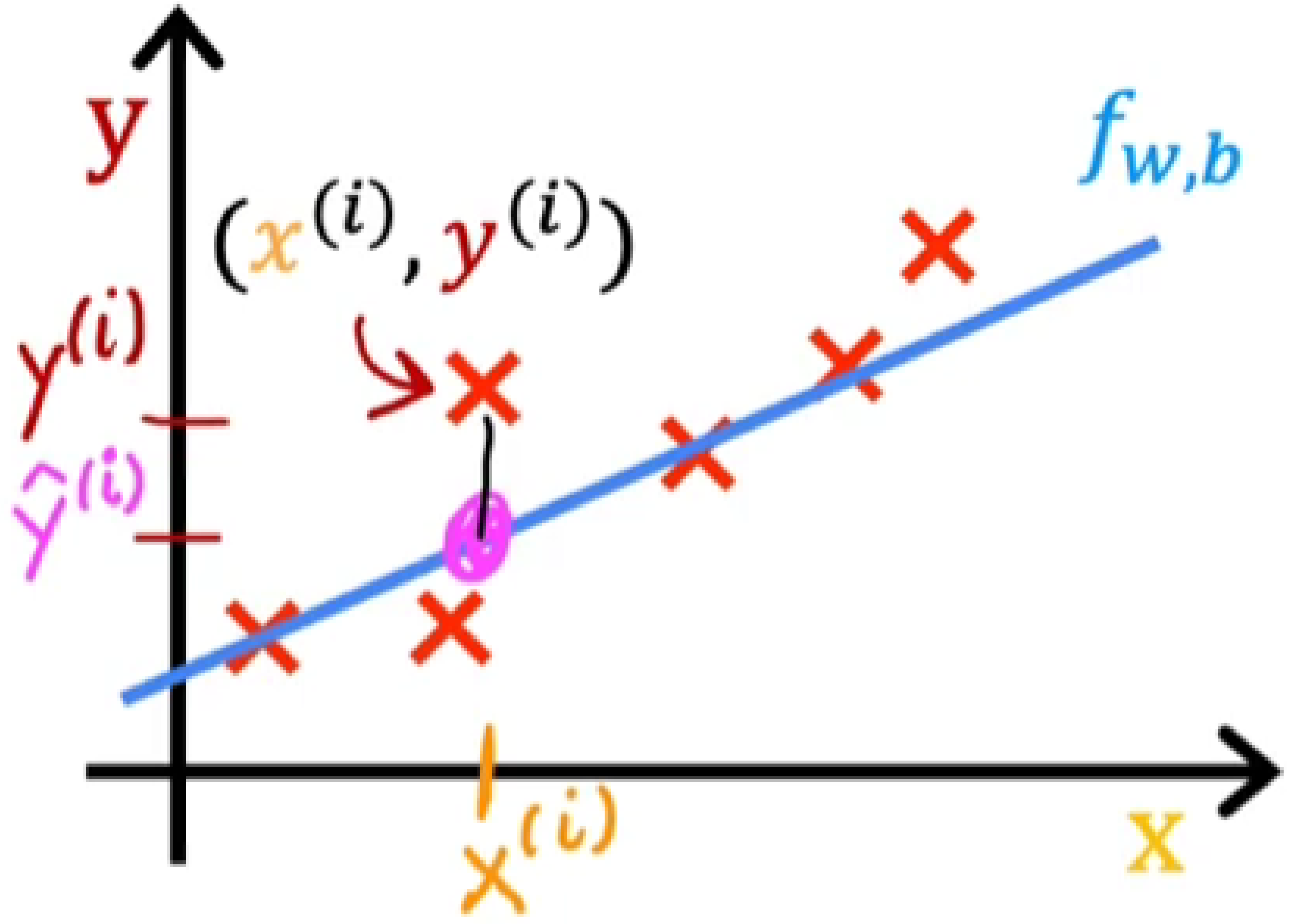
(x, y) = single training example

$(x^{(i)}, y^{(i)})$

$(x^{(i)}, y^{(i)})$ = i^{th} training example
(1^{st}, 2^{nd}, 3^{rd} ...)

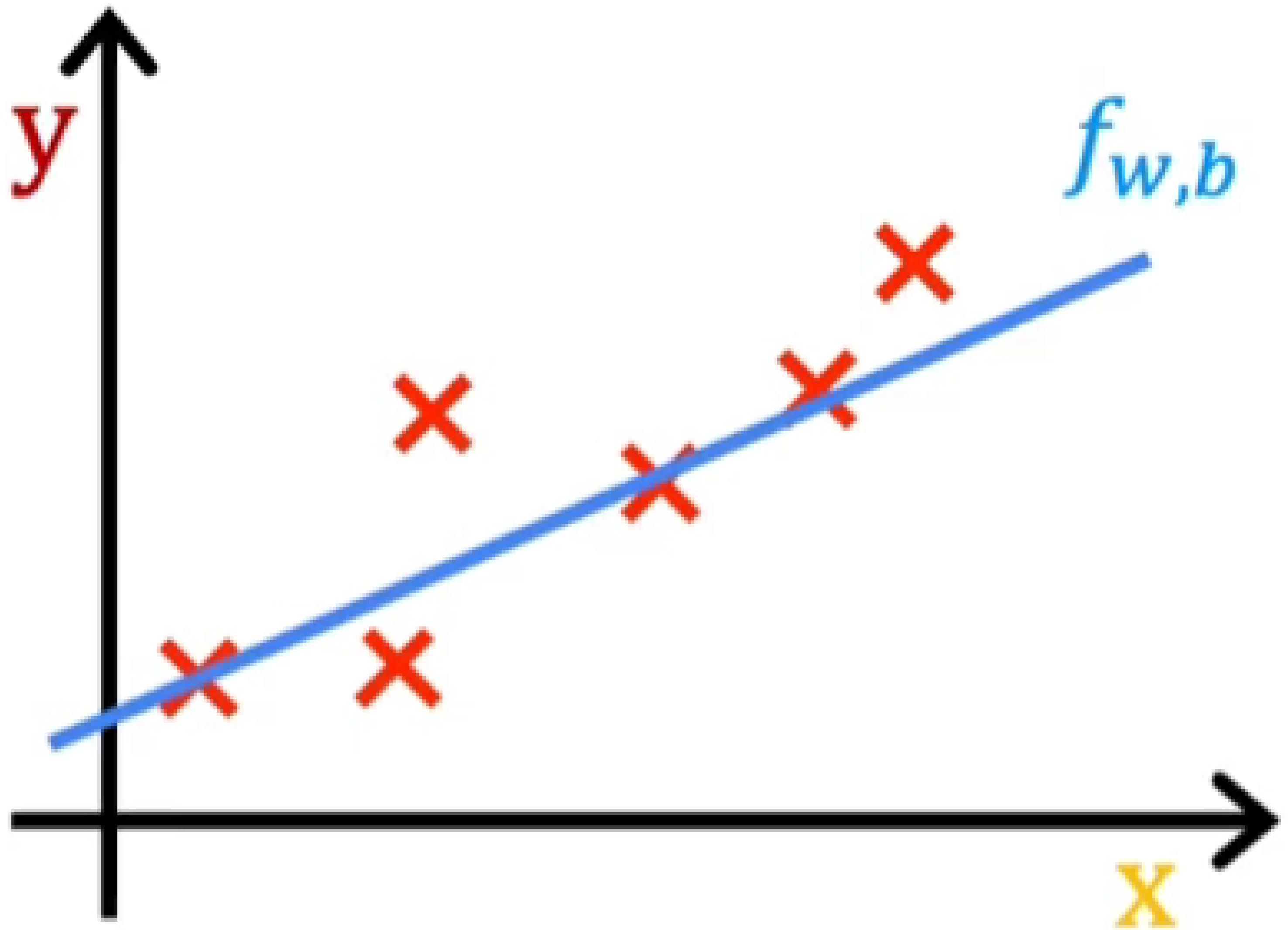
Error and Cost Function

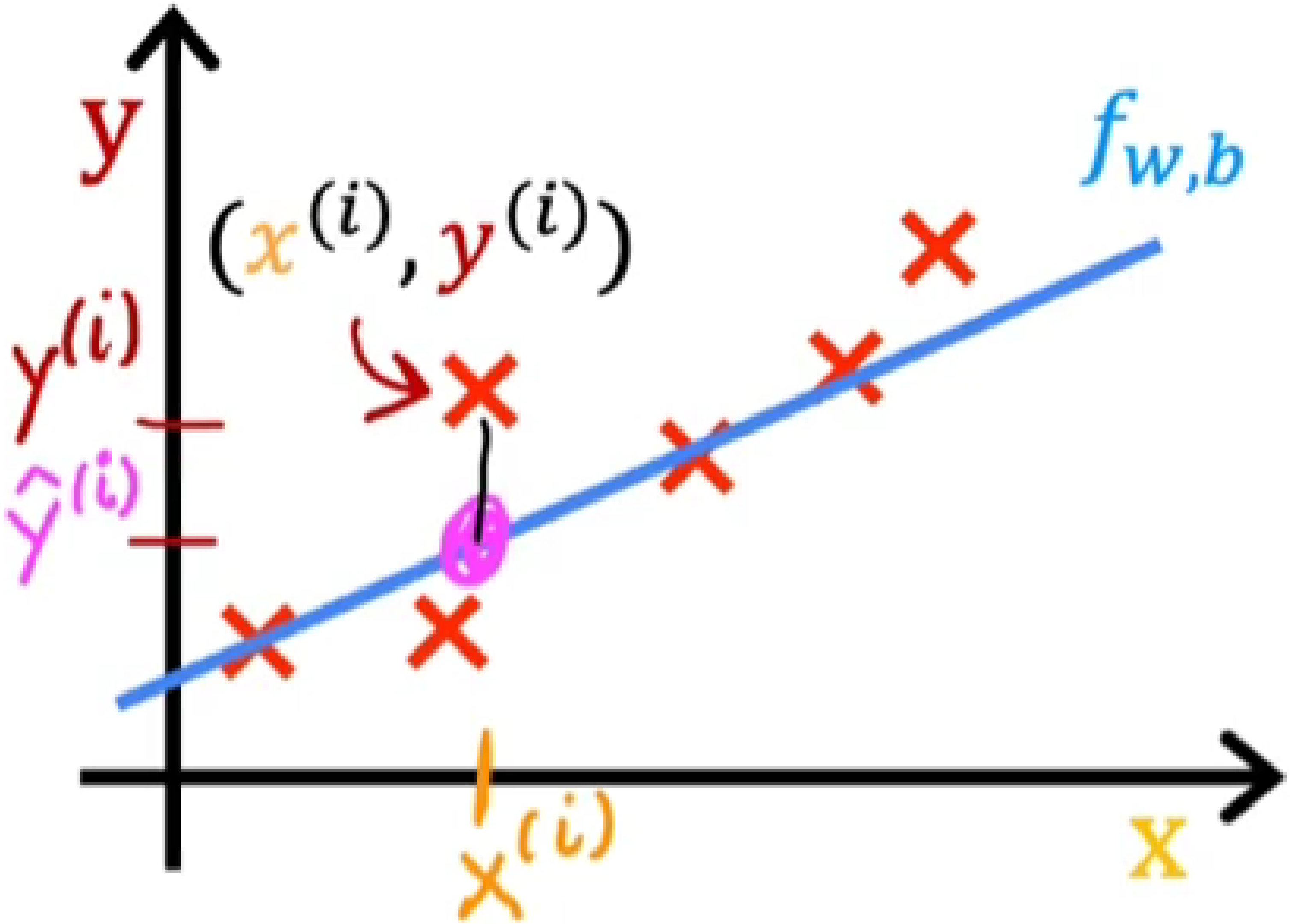




Error Function = $\left(\underbrace{\hat{y}^{(i)} - y^{(i)}}_{\text{error}} \right)^2$

Error and Cost Function



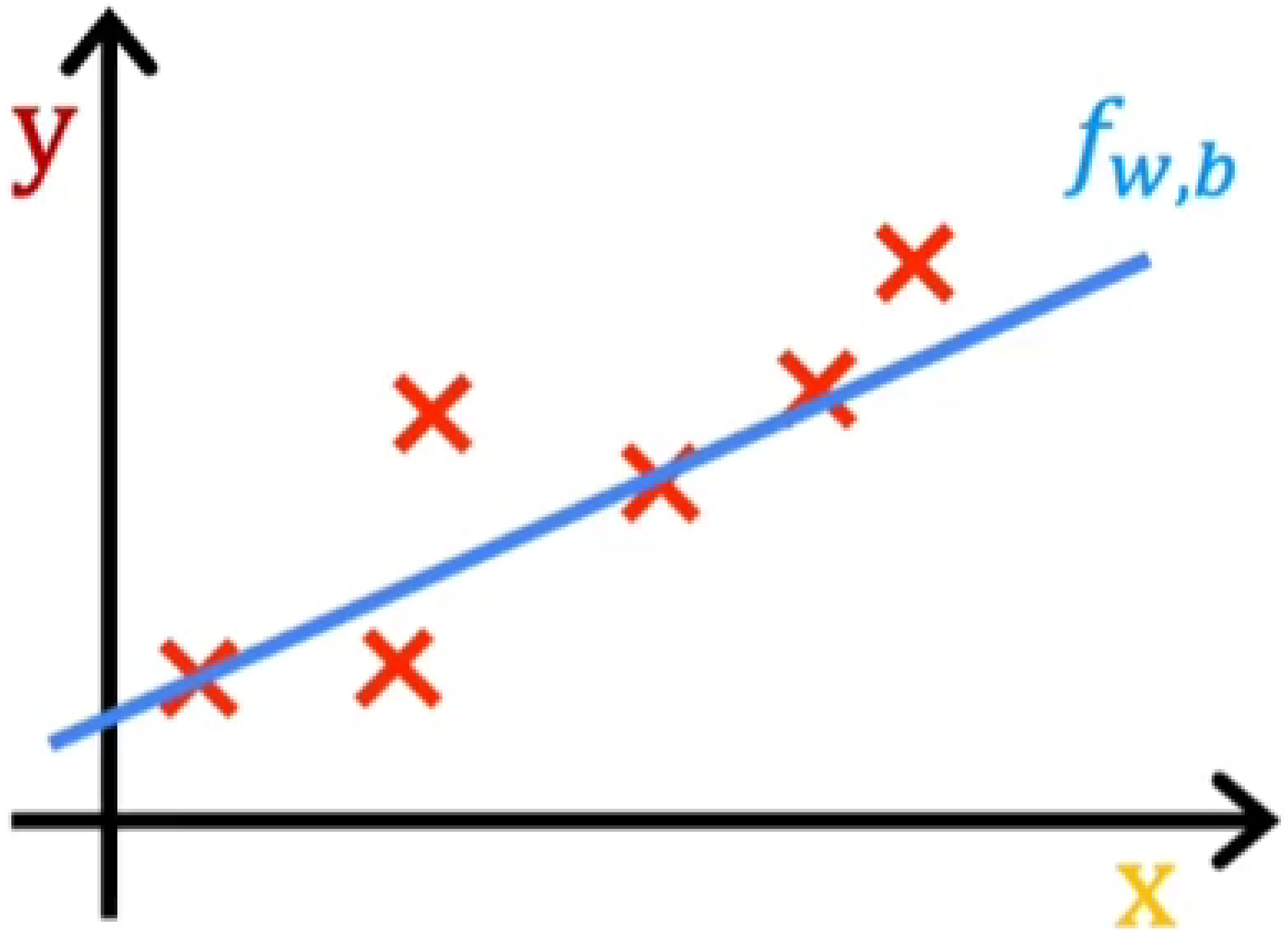


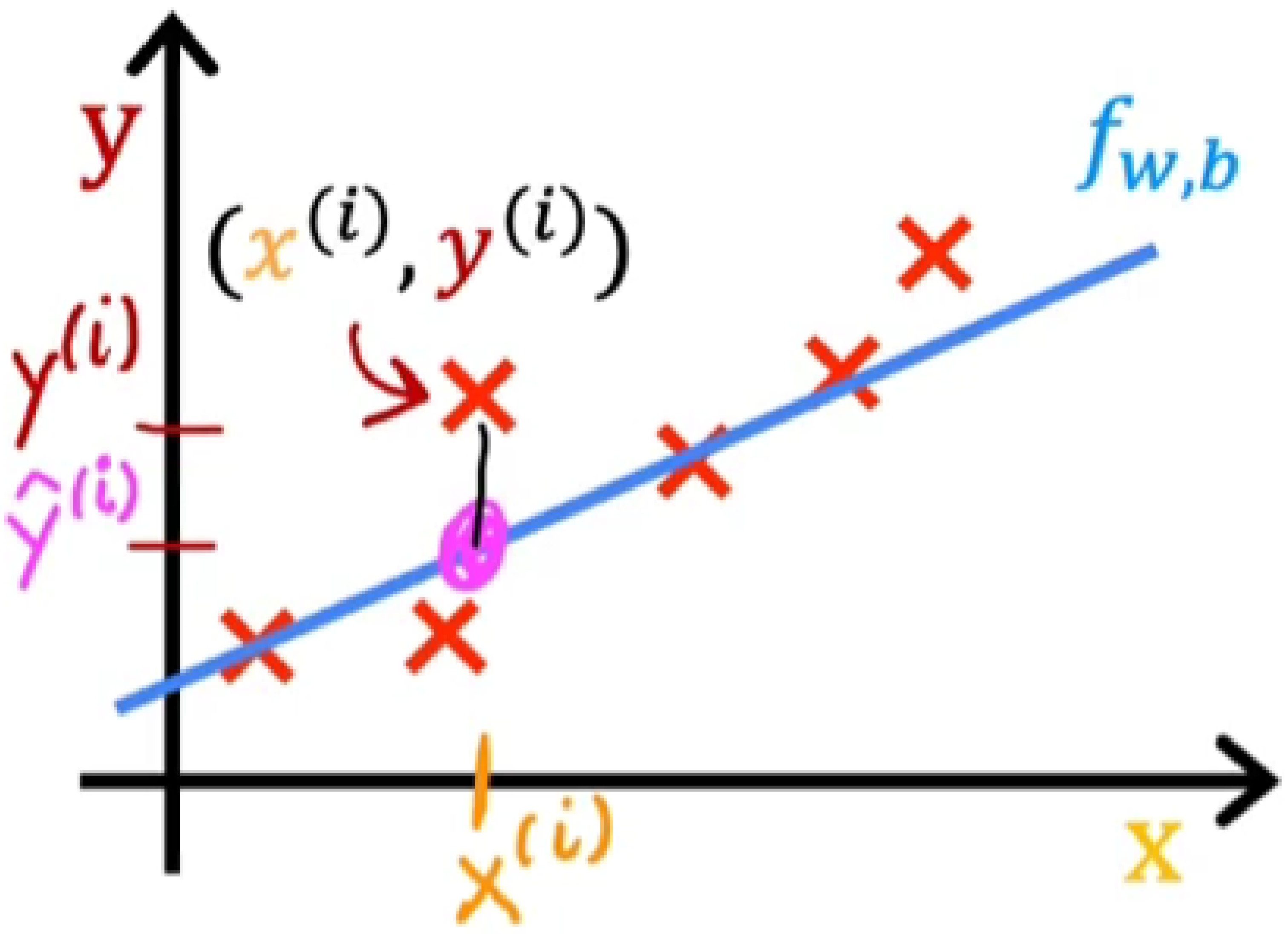
$$J(w,b) = \frac{1}{2m}\sum_{i=1}^{m}\left(\underbrace{\hat{y}^{(i)} - y^{(i)}}_{\text{error}}\right)^2$$

Cost Function

m = number of training examples

Error and Cost Function

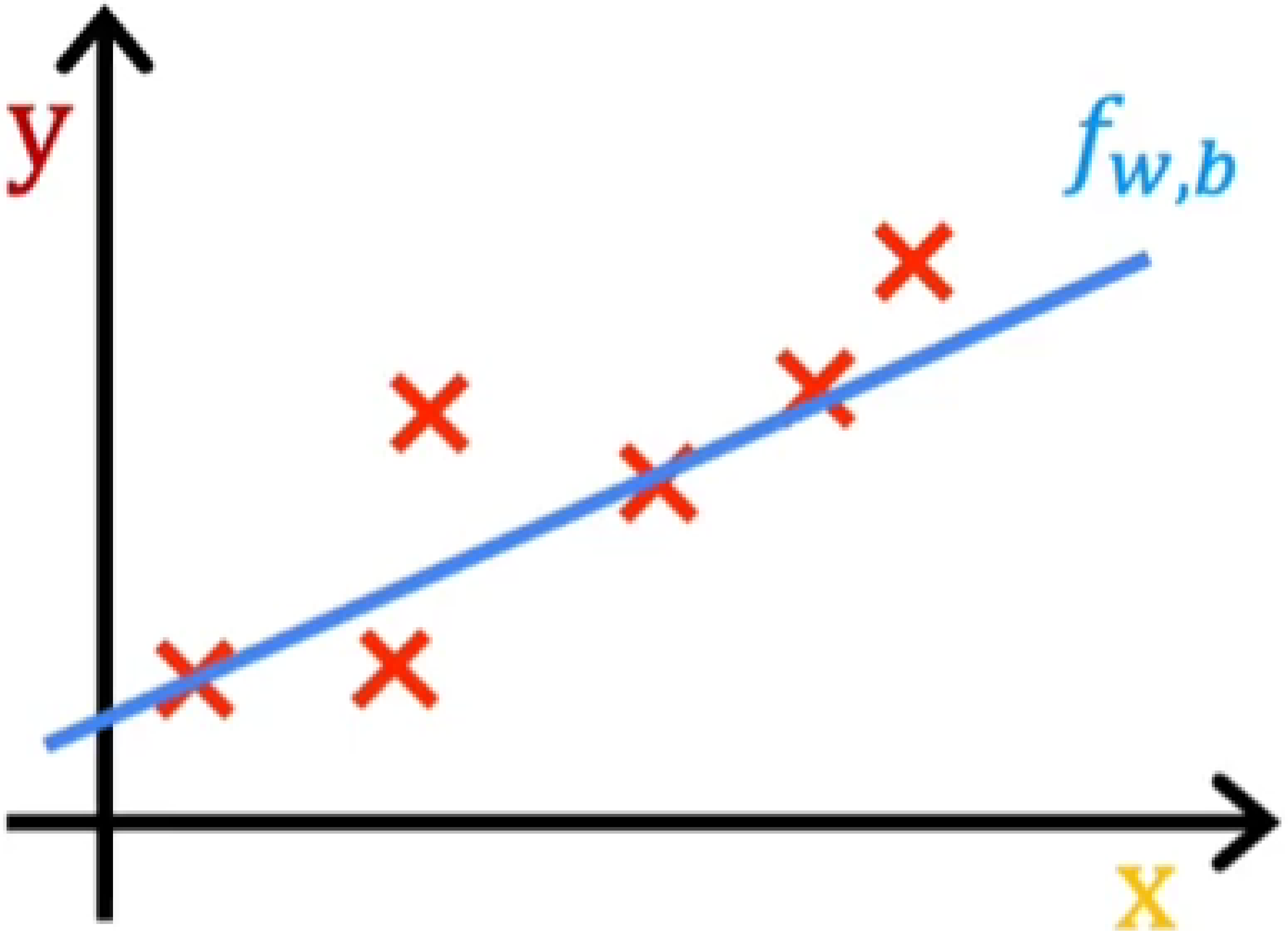


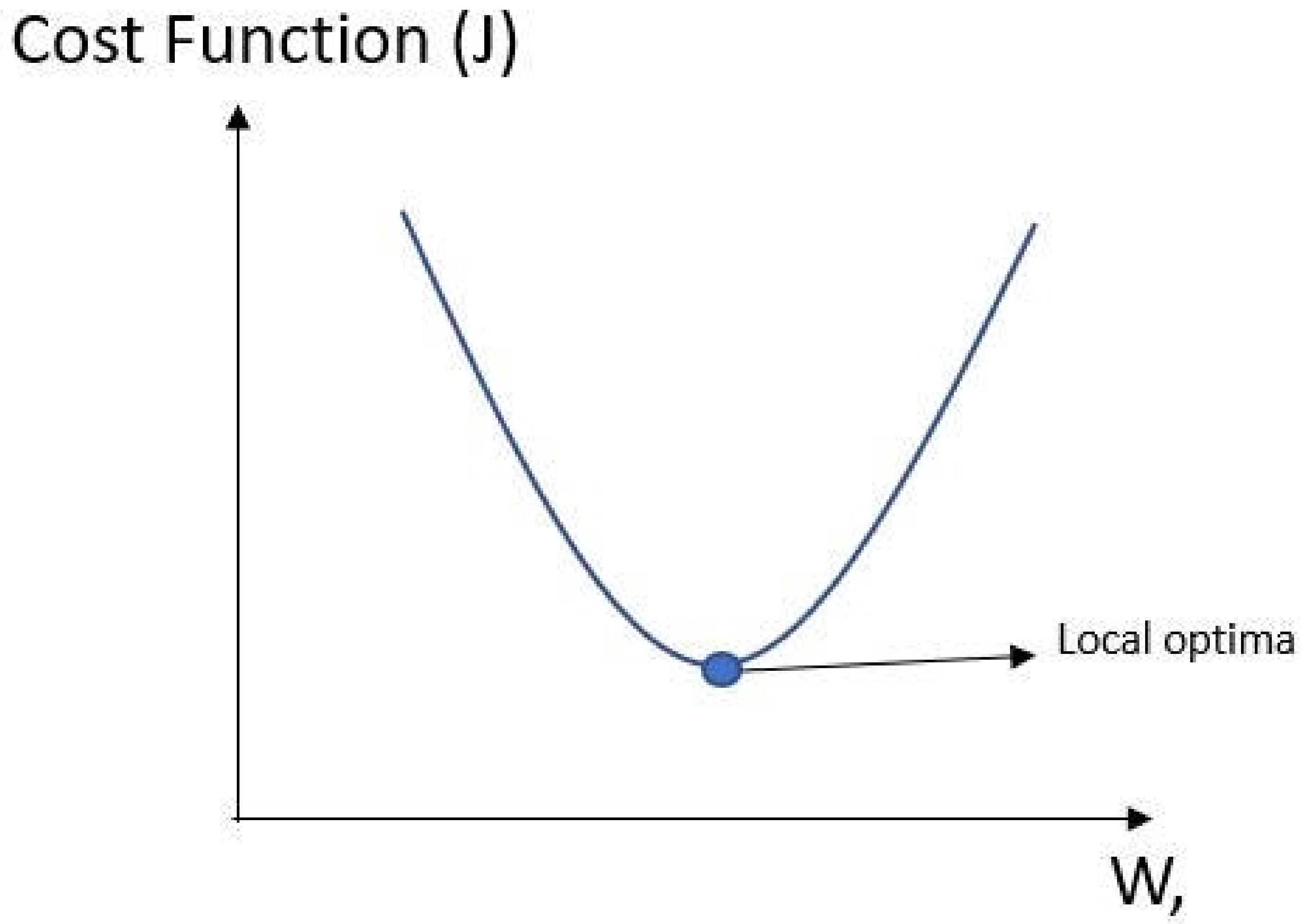


$$J(w,b) = \frac{1}{2m}\sum_{i=1}^{m}\left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right)^2$$

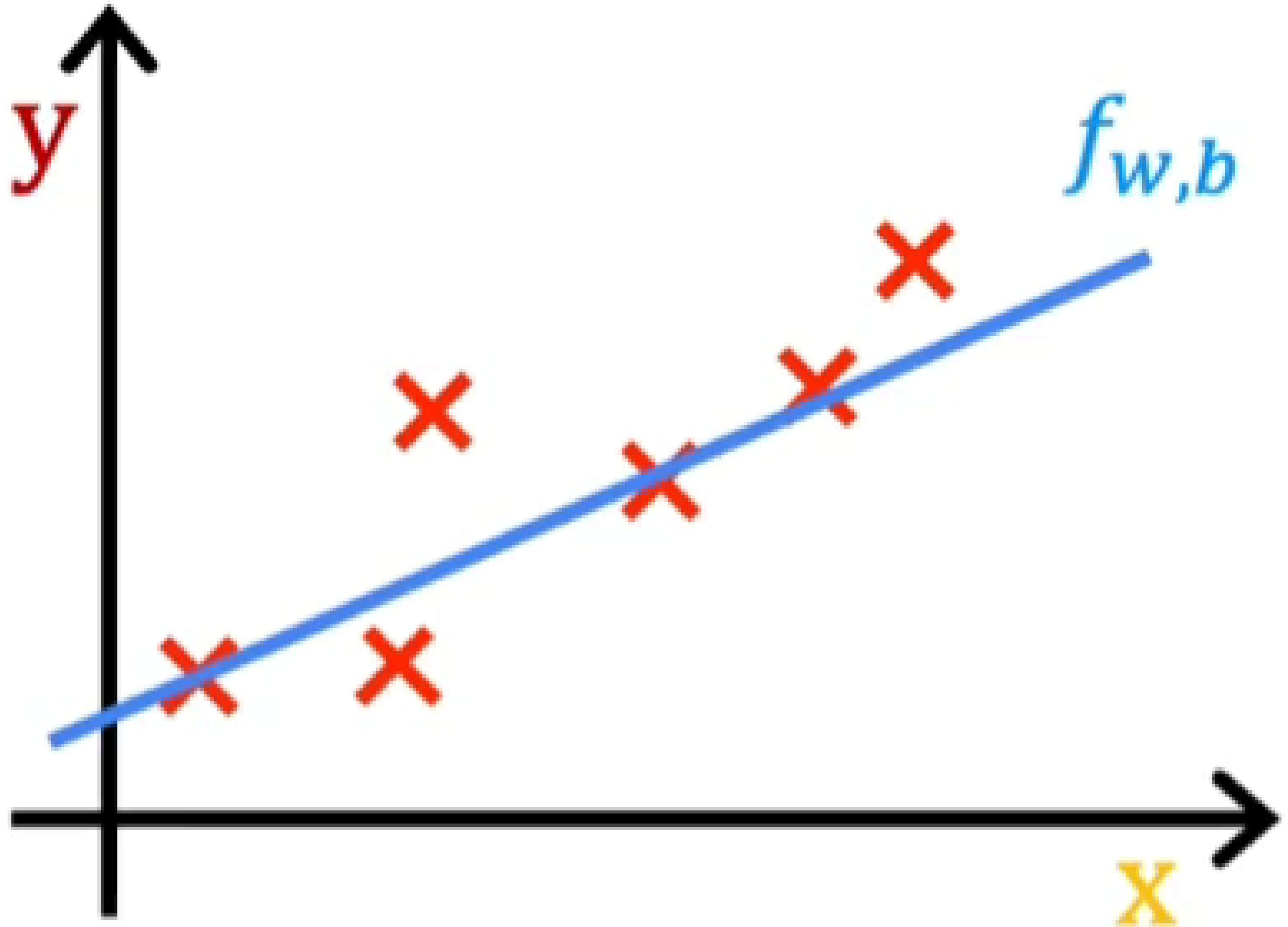
Cost Function

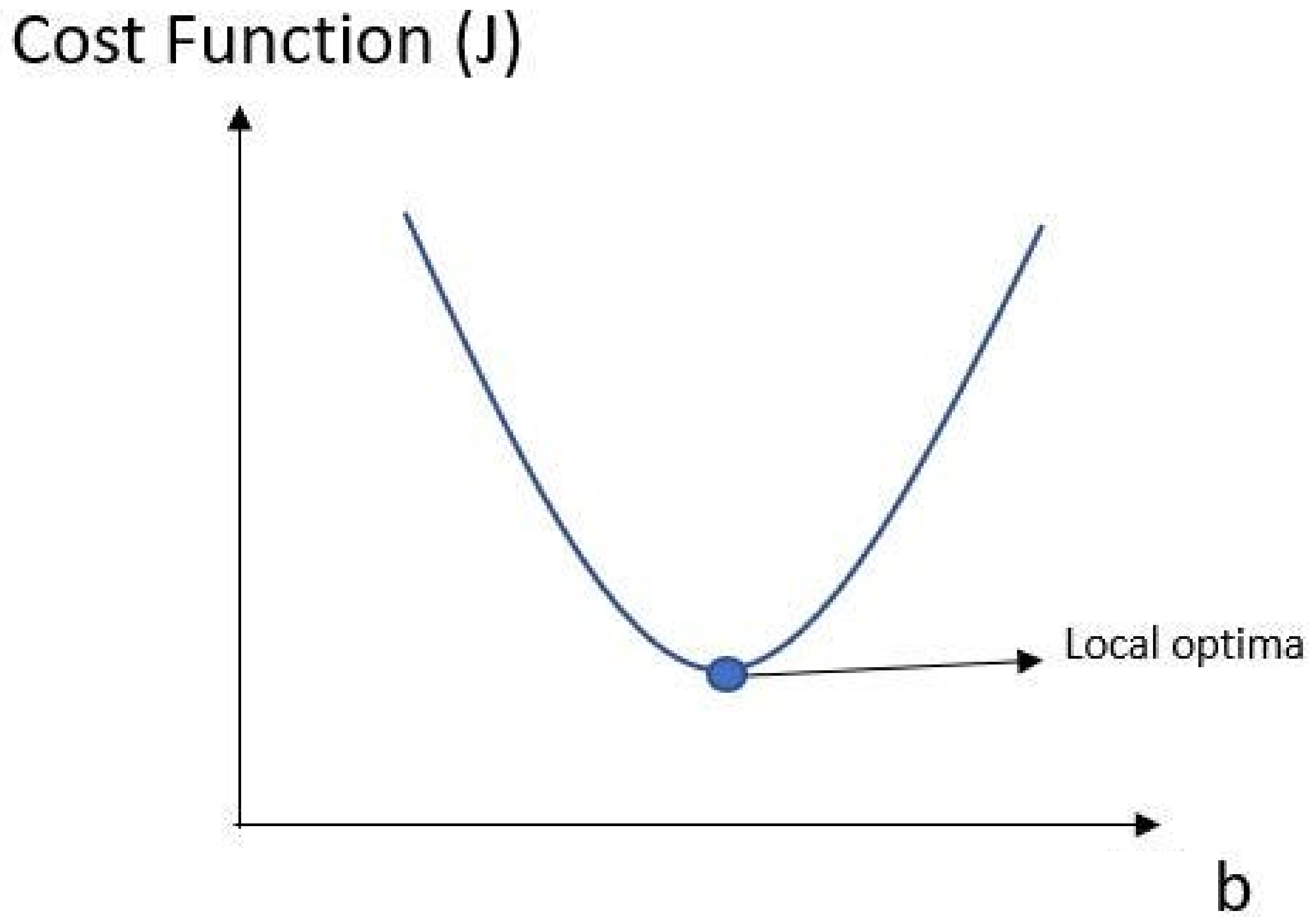
Error and Cost Function - Finding the sweet spot



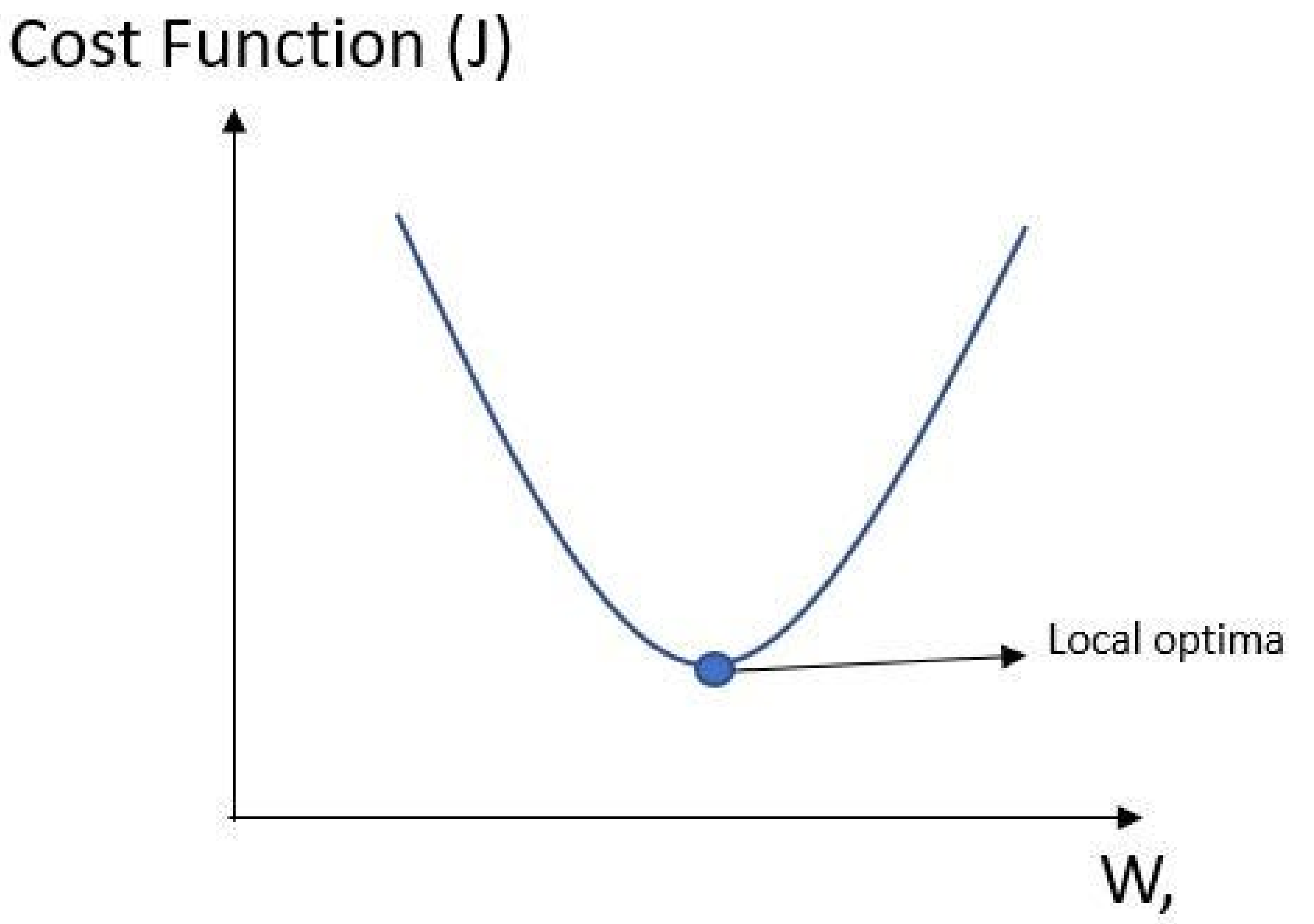


Error and Cost Function - Finding the sweet spot



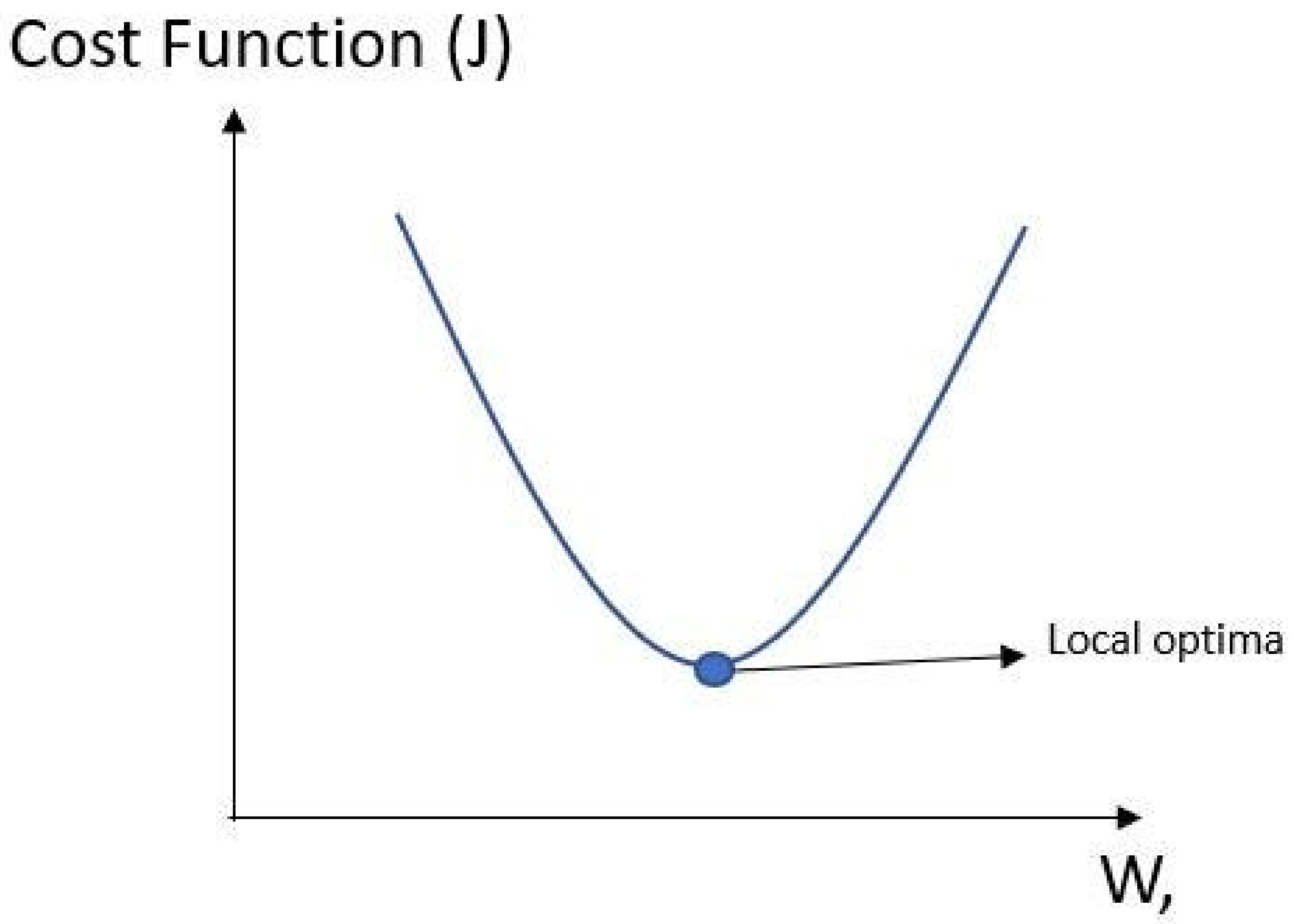


Gradient Descent



$$w = w - \alpha \frac{\partial}{\partial w} J(w,b)$$

Gradient Descent



$$w = w - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right) x^{(i)}$$

Gradient Descent - Derivation

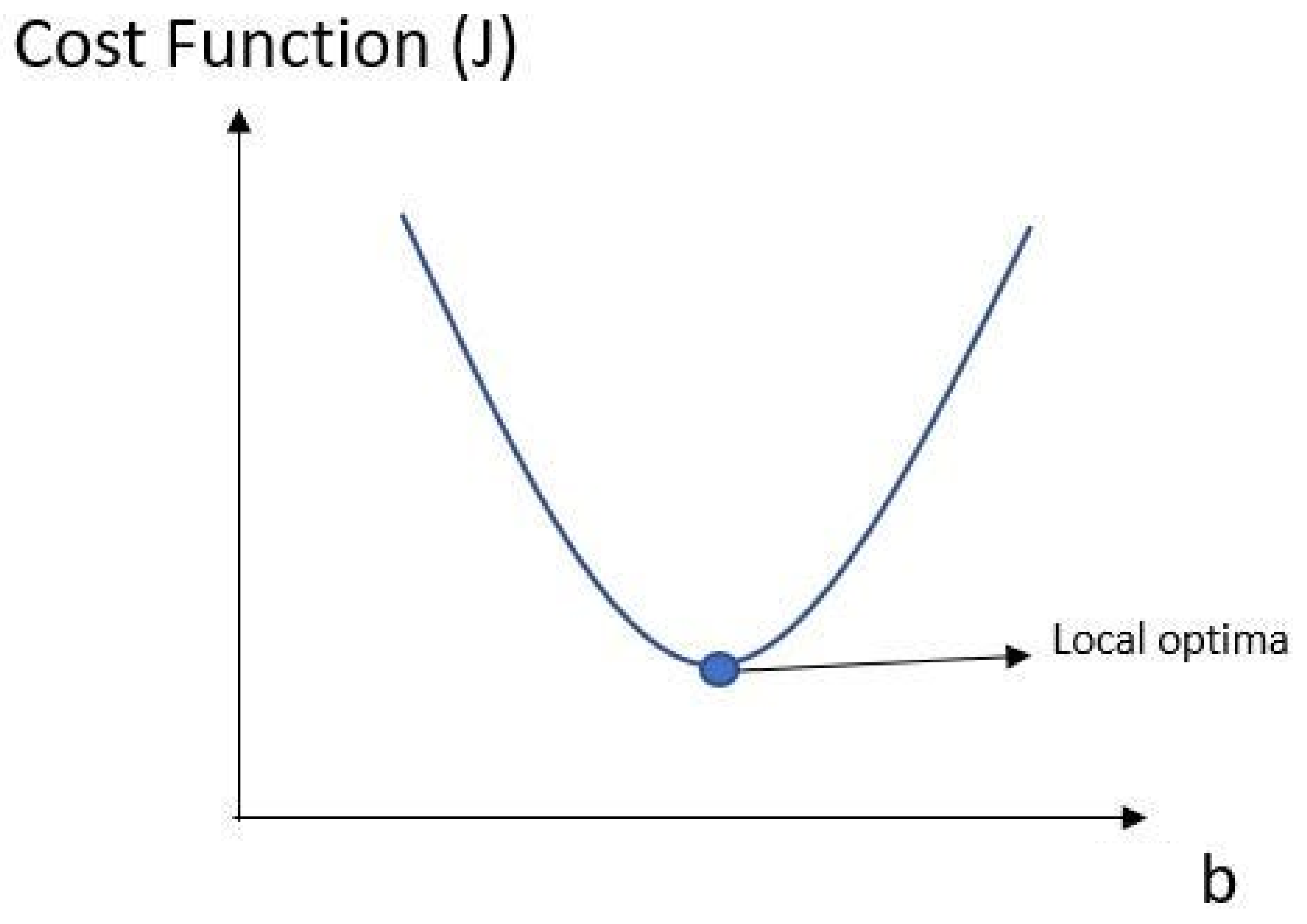
```
loss = (y_pred – y)²/n


loss = (y_pred² + y² – 2y*y_pred)/n
(expanding the whole square)


=>( (x*w+b)² + y² – 2y*(x*w+b))/n
(substitute y_pred)


=> ((x*w+b)²/n ) + (y²/n) +
((-2y(x*w+b))/n) (splitting the terms)


Let A = ((x*w+b)²/n )
Let B = (y²/n),
Let C = ((-2y(x*w+b))/n)
```

```
A = ( x²w² + b² + 2xwb )/n (expanding)
∂A/∂w = ( 2x²w + 2xb )/n (differentiating)


∂B/∂w = 0 (differentiating)


C = (-2yxw – 2yb)/n
∂C/∂w = (-2yx)/n (differentiating)
```

So, ∂loss/∂w will be the addition of all these terms:

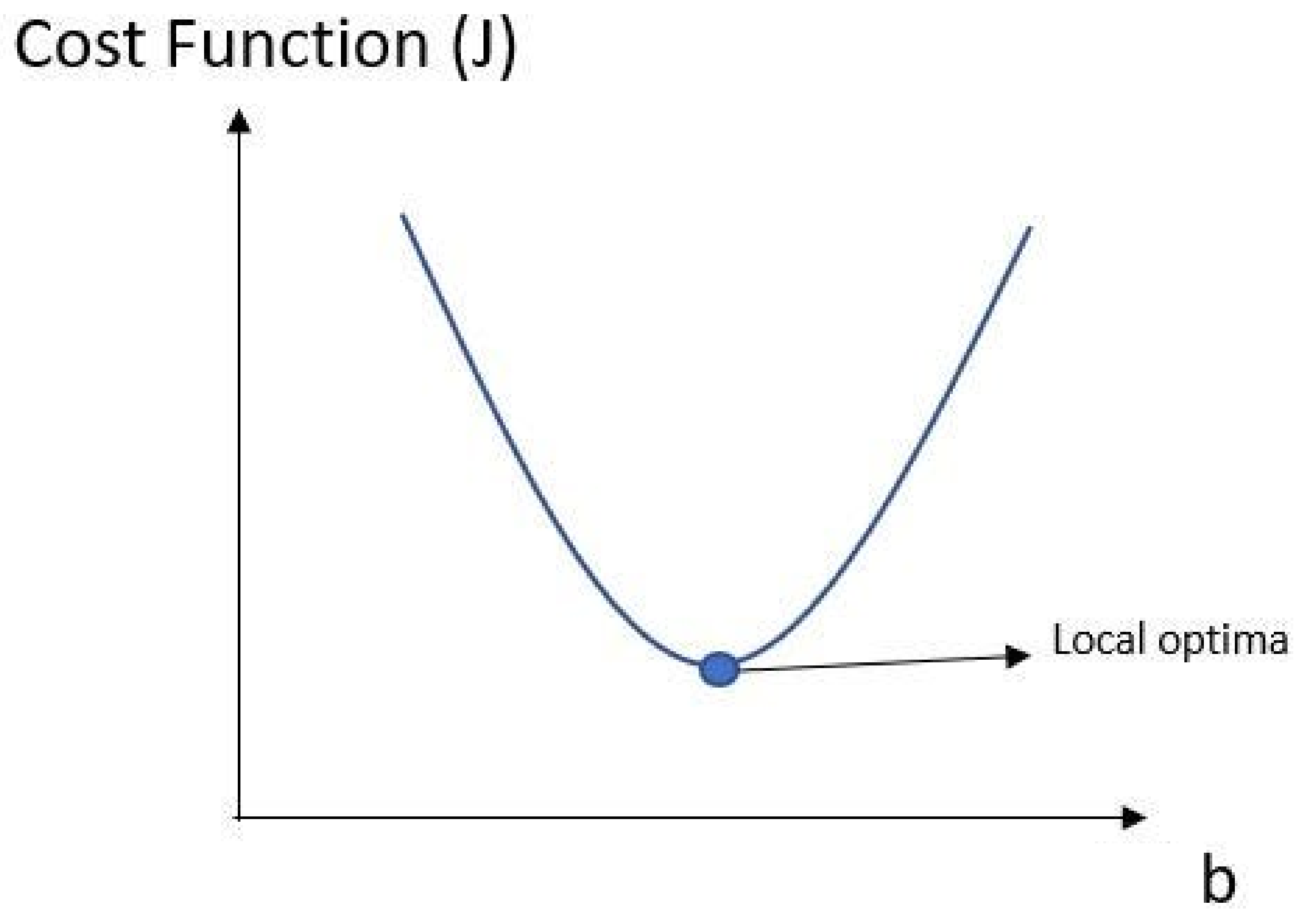
```
∂loss/∂w = (2x²w + wxb – 2yx)/n
=> (2x(x*w + b – y))/n
```

Gradient Descent



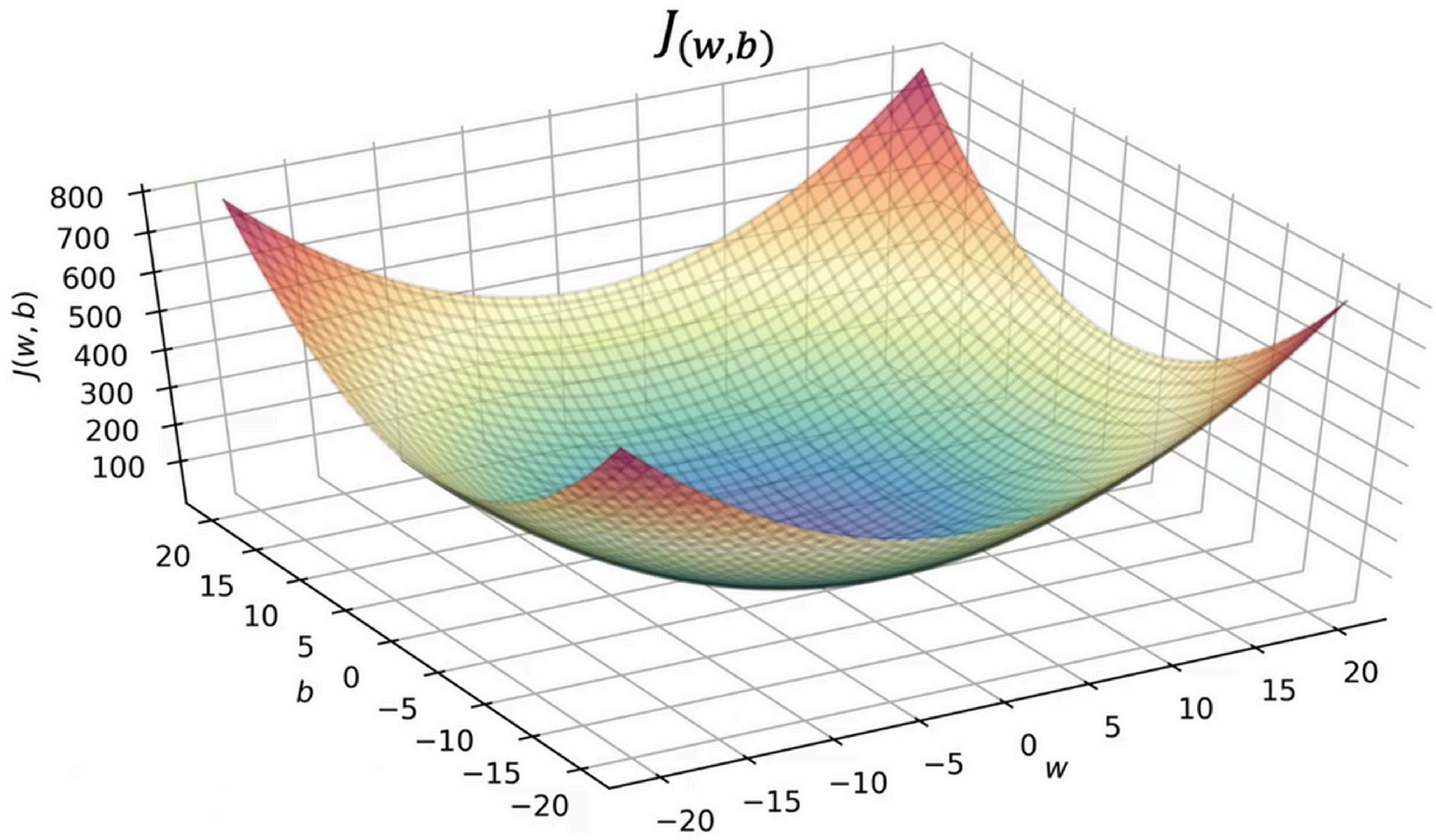
$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

Gradient Descent



$$b = b - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}\left(x^{(i)}\right) - y^{(i)} \right)$$

Gradient Descent



Multivariate Linear Regression

Size in feet^2 x_1	Number of bedrooms x_2	Number of floors x_3	Age of home in years x_4	Price ($) in $1000's
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
...	...	...	...	...

Multivariate Linear Regression

Size in feet^2	Number of bedrooms	Number of floors	Age of home in years	Price ($) in $1000's
x_1	x_2	x_3	x_4	
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
...	...	...	...	...

Previously: $f_{w,b}(x) = wx + b$

$$f_{w,b}(x) = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + b$$

example

$$f_{w,b}(x) = 0.1 \underset{\text{size}}{x_1} + 4 \underset{\#\text{bedrooms}}{x_2} + 10 \underset{\#\text{floors}}{x_3} + -2 \underset{\text{years}}{x_4} + \underset{\text{base price}}{80}$$

Multivariate Linear Regression

Previously: $f_{w,b}(x) = wx + b$

$$f_{w,b}(x) = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + b$$

example

$$f_{w,b}(x) = 0.1\, x_1 + 4\, x_2 + 10\, x_3 + -2\, x_4 + 80$$

size #bedrooms #floors years base price

Multivariate Linear Regression - Gradient Descent

$$w_1 = w_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)} \right) x_1^{(i)}$$

$$w_n = w_n - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)} \right) x_n^{(i)}$$

$$b = b - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)} \right)$$