$f_{w,b}(size)$

starting point

$J(w,b)$

price in $1000's

Cost: 77237
Cost: 45401
Cost: 29109
Cost: 19076
Cost: 13318
Cost: 9460
Cost: 7209
Cost: 3559
Cost: 2311

we reach the middle point after applying multiple iterations of gradient descent

best fitting line

size in feet²

$J(w,b)$
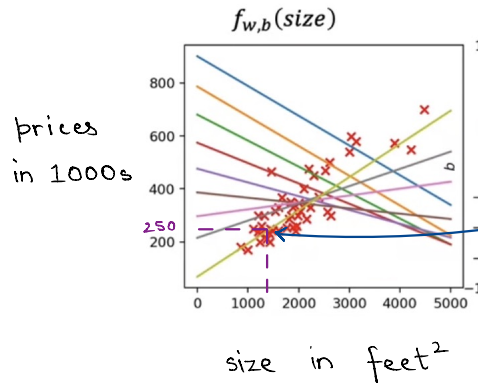
Now, we can use this $f_{w,b}(x)$ model to predict other house prices.

suppose house A is 1250 sq. ft. then price will be $250,000.

$f_{w,b}(size)$

prices in 1000s

250
200

size in feet²

To be precise, this type of algorithm is called
batch gradient descent (there are other types
such as subset gradient descent, etc)

Batch gradient descent uses all of the
training examples at each step.

x size in feet^2	y price in \$1000's
(1) 2104	400
(2) 1416	232
(3) 1534	315
(4) 852	178
... ...	...
(47) 3210	870

all examples will be tested for
every w, b.

$$m = 47 \rightarrow \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right)^2$$

Batch gradient descent is the most used
algorithm for linear regression.