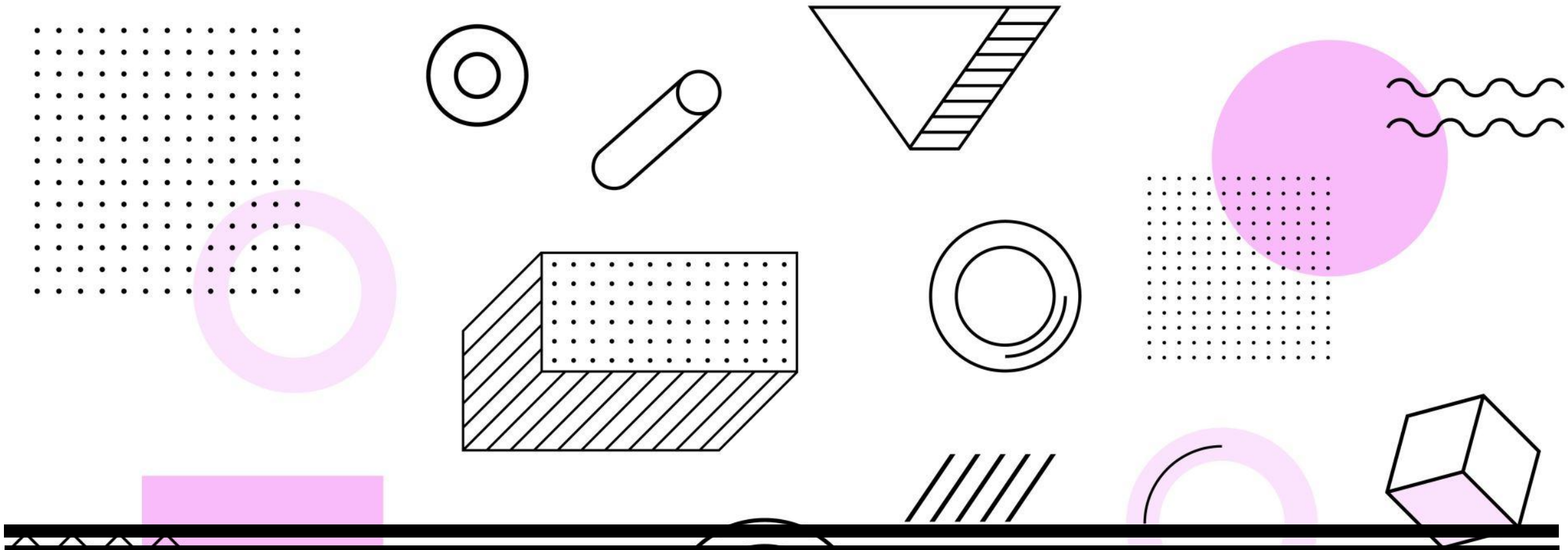


FEATURE SCALING

FEATURE AND PARAMETER VALUES

$b = 0$ { simplicity }

$w_1 x_1 + w_2 x_2 + b$

$x_1 = 2$; $x_2 = 35000$; $y = 50$ ✓

$w_1 = 100$, $w_2 = 100$

$\hat{y} = 35002000$ ✓

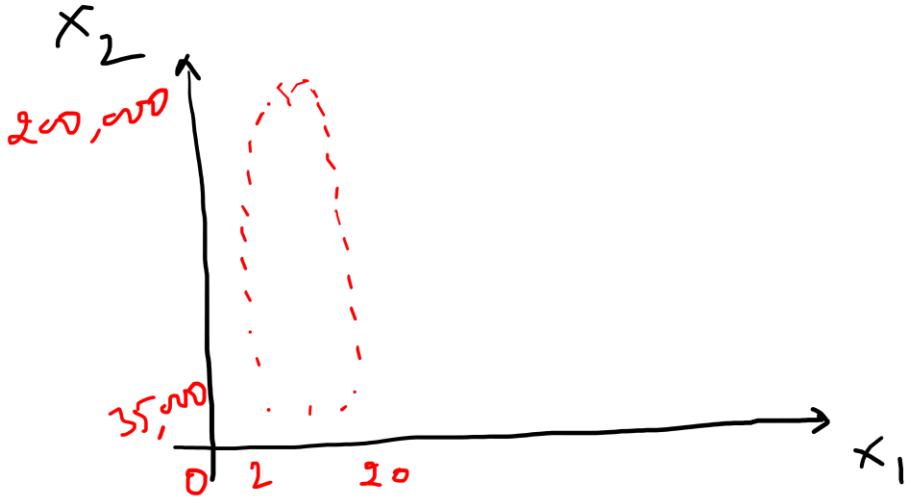
✗

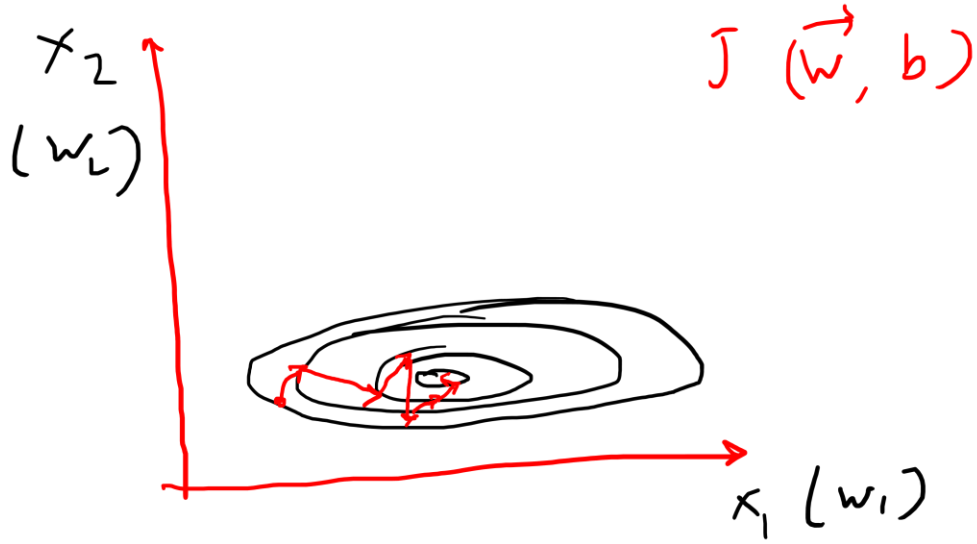
$w_1 = 10$, $w_2 = 0.001$

$\hat{y} = 55$ ✓

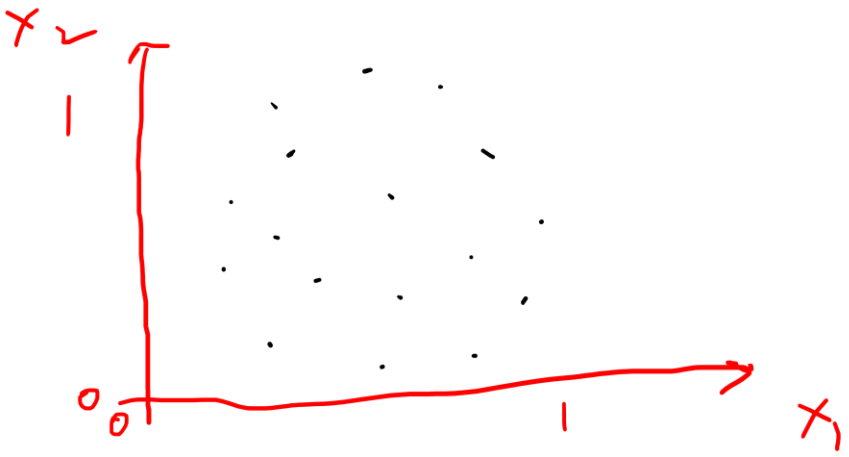
FEATURE SIZE AND PARAMETER SIZE

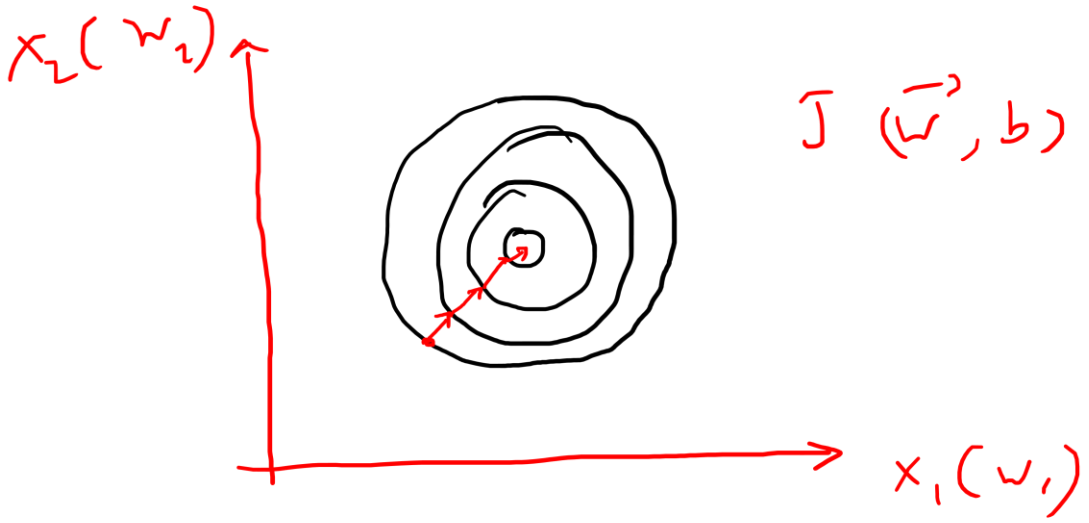
	Size of feature (x_j)	Size of parameter (w_j)
x_1 (sqft)	Small	Large
x_2 (Scalarg)	Large	Small





FEATURE SIZE AND GRADIENT DESCENT





FEATURE SCALING

- Feature range is either very large or very small

FEATURE SCALING

- Division by Max number in the feature
- Mean Normalization
- Z-Score Normalization

DATASET

Performance Score (Target Variable): Represents an employee's expected performance on a **scale of 0 to 100**, based on past data.

Experience (Years)	Salary ($)	Performance Score (Target)
2	35,000	50
5	50,000	55
7	80,000	70
3	45,000	52
10	120,000	85
15	200,000	95
20	150,000	92
12	90,000	80
6	60,000	65
18	170,000	90

DIVISION BY MAX NUMBER IN THE FEATURE (MIN-MAX SCALING)

$35{,}000 \leq x_2 \leq 2{,}00{,}000$

$2 \leq x_1 \leq 20$

$x_2 = 45{,}000$

$x_{2,s} = \frac{45000}{200000} = 0.225$

$x_1 = 3$

$x_{1,s} = \frac{3}{20} = 0.15$

$0.175 \leq x_{2,s} \leq 1$

$0.1 \leq x_{1,s} \leq 1$

MEAN NORMALIZATION

$\mu_1 = 9.8$

$x_1 = 3$

$$x_{1,s} = \frac{x_1 - \mu_1}{Max - Min}$$

$x_{1,s} = -0.37$

$-0.43 \leq x_{1,s} \leq 0.56$

$\mu_2 = 1,00,000$

$x_2 = 45,000$

$$x_{2,s} = \frac{x_2 - \mu_2}{Max - Min}$$

$x_{2,s} = -0.33$

$-0.33 \leq x_{2,s} \leq 0.60$

Z-SCORE NORMALIZATION

σ_1 = ? Std. Deviation (x_1)

σ_2 = ? Std. Deviation (x_2)

μ_1 = Avg (x_1)

μ_2 = Avg (x_2)

$$x_{1,s} = \frac{x_1 - \mu_1}{\sigma_1} \quad ; \quad x_{2,s} = \frac{x_2 - \mu_2}{\sigma_2}$$