



# Regresja

Weronika Piotrowska

Golem 2023

# Agenda

## I. Część teoretyczna


1. Wprowadzenie
2. Regresja czy klasyfikacja?
3. Prosty przykład
4. Mean Squared Error
5. Spadek gradientu
6. Trochę mniej prosty przykład
7. Regularyzacja
8. Czeka, dlaczego my to w ogóle robimy?
9. Regresja wielomianowa

## II. Część praktyczna

# Przypomnienie

Na podstawie tego

Chcemy przewidzieć to



No	X1 transaction date	X2 house age	X3 distance to the nearest MRT station	X4 number of convenience stores	X5 latitude	X6 longitude	Y house price of unit area
1	2012.917	32	84.87882	10	24.98298	121.54024	37.9
2	2012.917	19.5	306.5947	9	24.98034	121.53951	42.2
3	2013.583	13.3	561.9845	5	24.98746	121.54391	47.3
4	2013.5	13.3	561.9845	5	24.98746	121.54391	54.8
5	2012.833	5	390.5684	5	24.97937	121.54245	43.1
6	2012.667	7.1	2175.03	3	24.96305	121.51254	32.1
7	2012.667	34.5	623.4731	7	24.97933	121.53642	40.3
8	2013.417	20.3	287.6025	6	24.98042	121.54228	46.7
9	2013.5	31.7	5512.038	1	24.95095	121.48458	18.8
10	2013.417	17.9	1783.18	3	24.96731	121.51486	22.1
11	2013.083	34.8	405.2134	1	24.97349	121.53372	41.4
12	2013.333	6.3	90.45606	9	24.97433	121.5431	58.1
13	2012.917	13	492.2313	5	24.96515	121.53737	39.3
14	2012.667	20.4	2469.645	4	24.96108	121.51046	23.8
15	2013.5	13.2	1164.838	4	24.99156	121.53406	34.3

# Przypomnienie

features

target / label



A diagram above the table shows a blue bracket spanning columns 1 through 7, labeled 'features', and a green bracket spanning column 8, labeled 'target / label'.

No	X1 transaction date	X2 house age	X3 distance to the nearest MRT station	X4 number of convenience stores	X5 latitude	X6 longitude	Y house price of unit area
1	2012.917	32	84.87882	10	24.98298	121.54024	37.9
2	2012.917	19.5	306.5947	9	24.98034	121.53951	42.2
3	2013.583	13.3	561.9845	5	24.98746	121.54391	47.3
4	2013.5	13.3	561.9845	5	24.98746	121.54391	54.8
5	2012.833	5	390.5684	5	24.97937	121.54245	43.1
6	2012.667	7.1	2175.03	3	24.96305	121.51254	32.1
7	2012.667	34.5	623.4731	7	24.97933	121.53642	40.3
8	2013.417	20.3	287.6025	6	24.98042	121.54228	46.7
9	2013.5	31.7	5512.038	1	24.95095	121.48458	18.8
10	2013.417	17.9	1783.18	3	24.96731	121.51486	22.1
11	2013.083	34.8	405.2134	1	24.97349	121.53372	41.4
12	2013.333	6.3	90.45606	9	24.97433	121.5431	58.1
13	2012.917	13	492.2313	5	24.96515	121.53737	39.3
14	2012.667	20.4	2469.645	4	24.96108	121.51046	23.8
15	2013.5	13.2	1164.838	4	24.99156	121.53406	34.3

# Regresja vs klasyfikacja

## Regresja - **liczby**

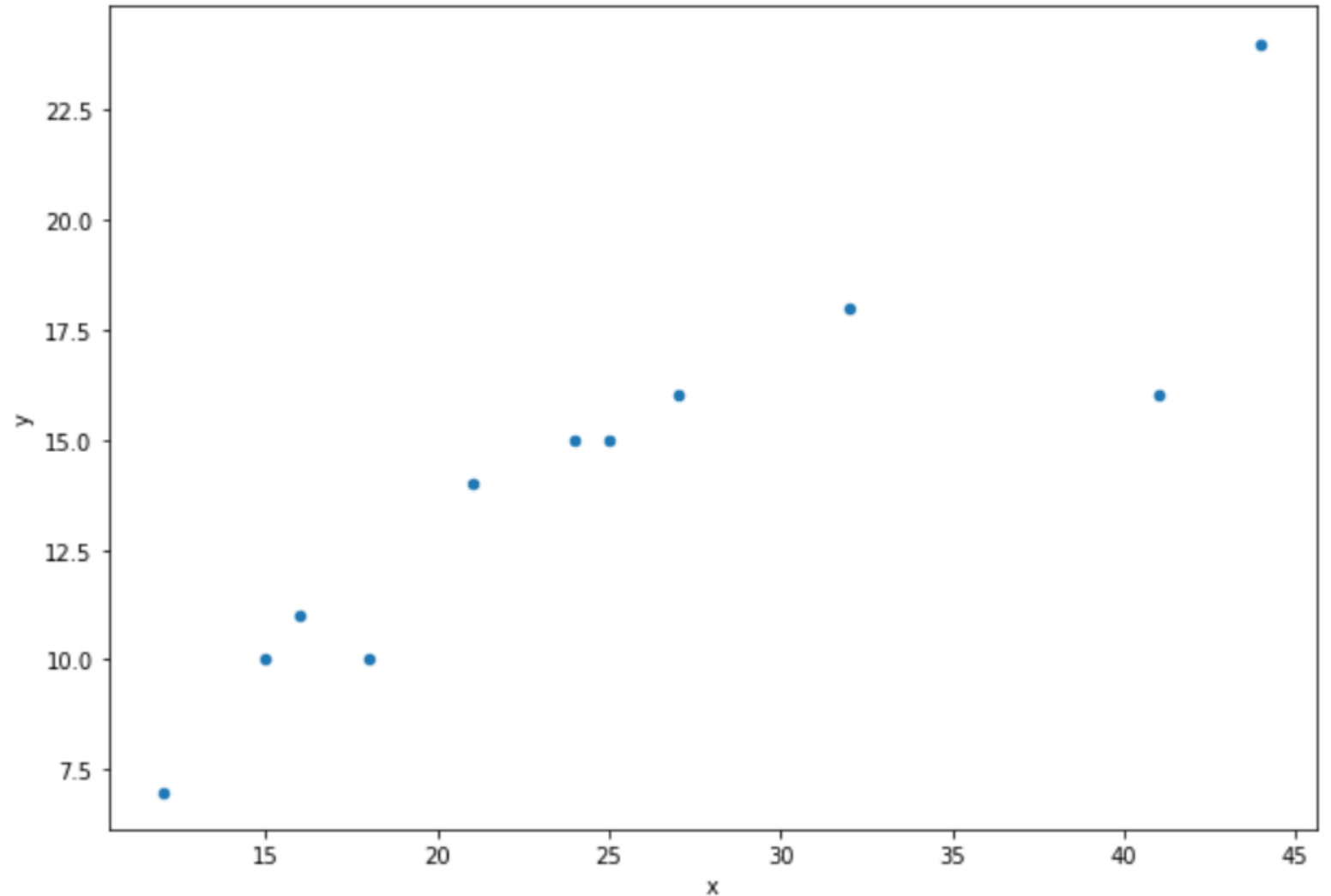
- wartości funkcji
- cena
- ilość ludzi
- jakość (np w skali 1-5)

## Klasyfikacja - **kategorie**

- tak/nie
- rasy psów
- grupa krwi
- jaki obiekt?

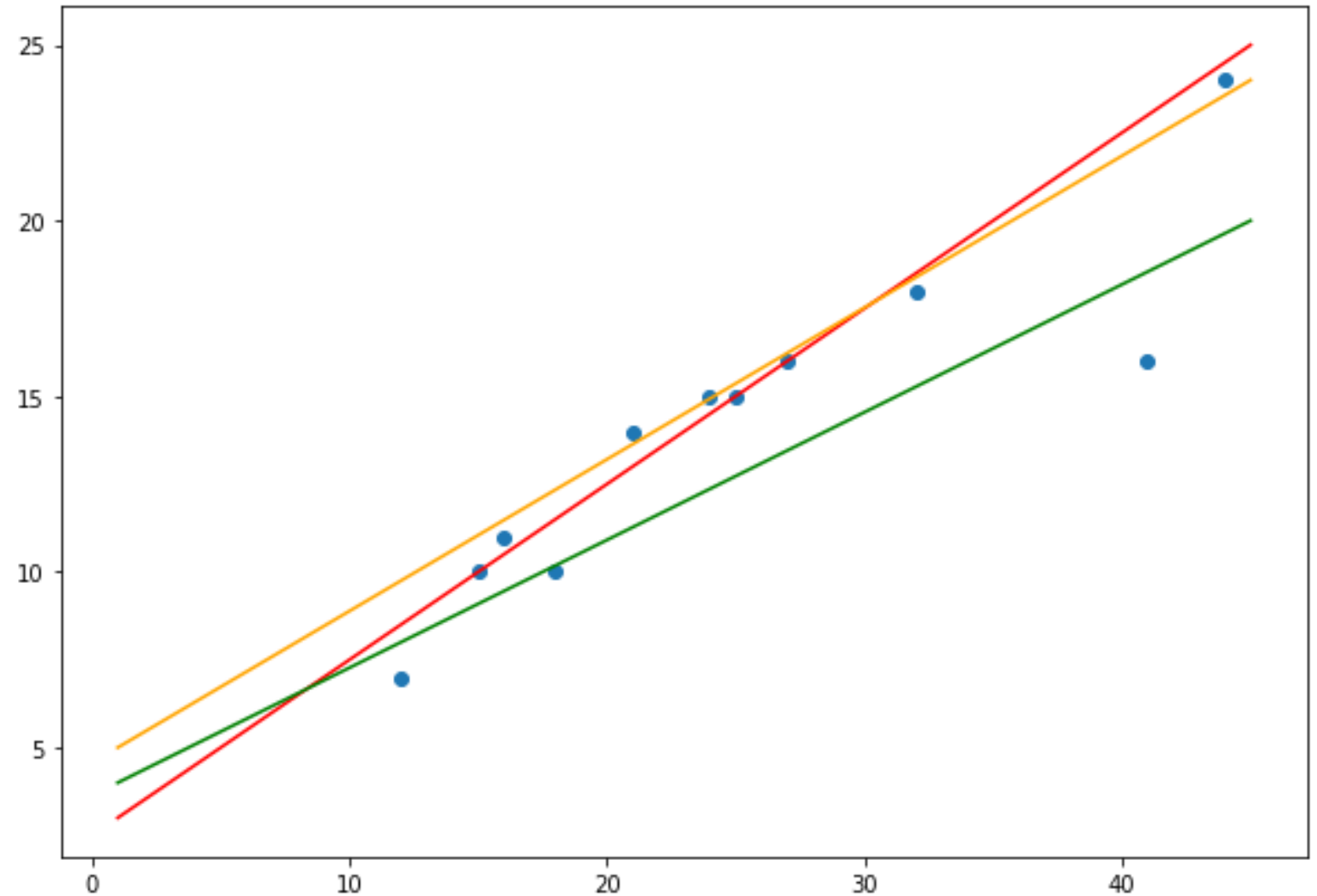
# Najprostszy przypadek

	x	y
0	12	7
1	15	10
2	16	11
3	18	10
4	21	14
5	24	15
6	25	15
7	27	16
8	32	18
9	41	16
10	44	24



# Najprostszy przypadek

	x	y
0	12	7
1	15	10
2	16	11
3	18	10
4	21	14
5	24	15
6	25	15
7	27	16
8	32	18
9	41	16
10	44	24



# Czego tak właściwie szukamy?

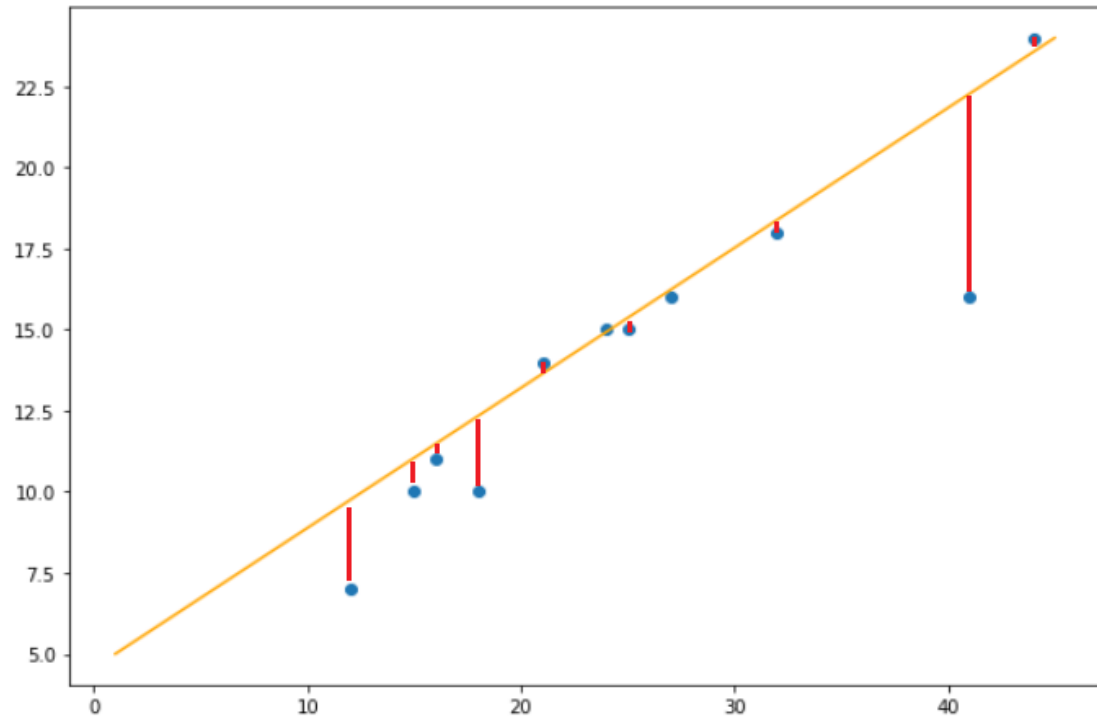
Najlepszego  $\beta_1, \beta_0$  dla

$$y = \beta_1 x + \beta_0$$



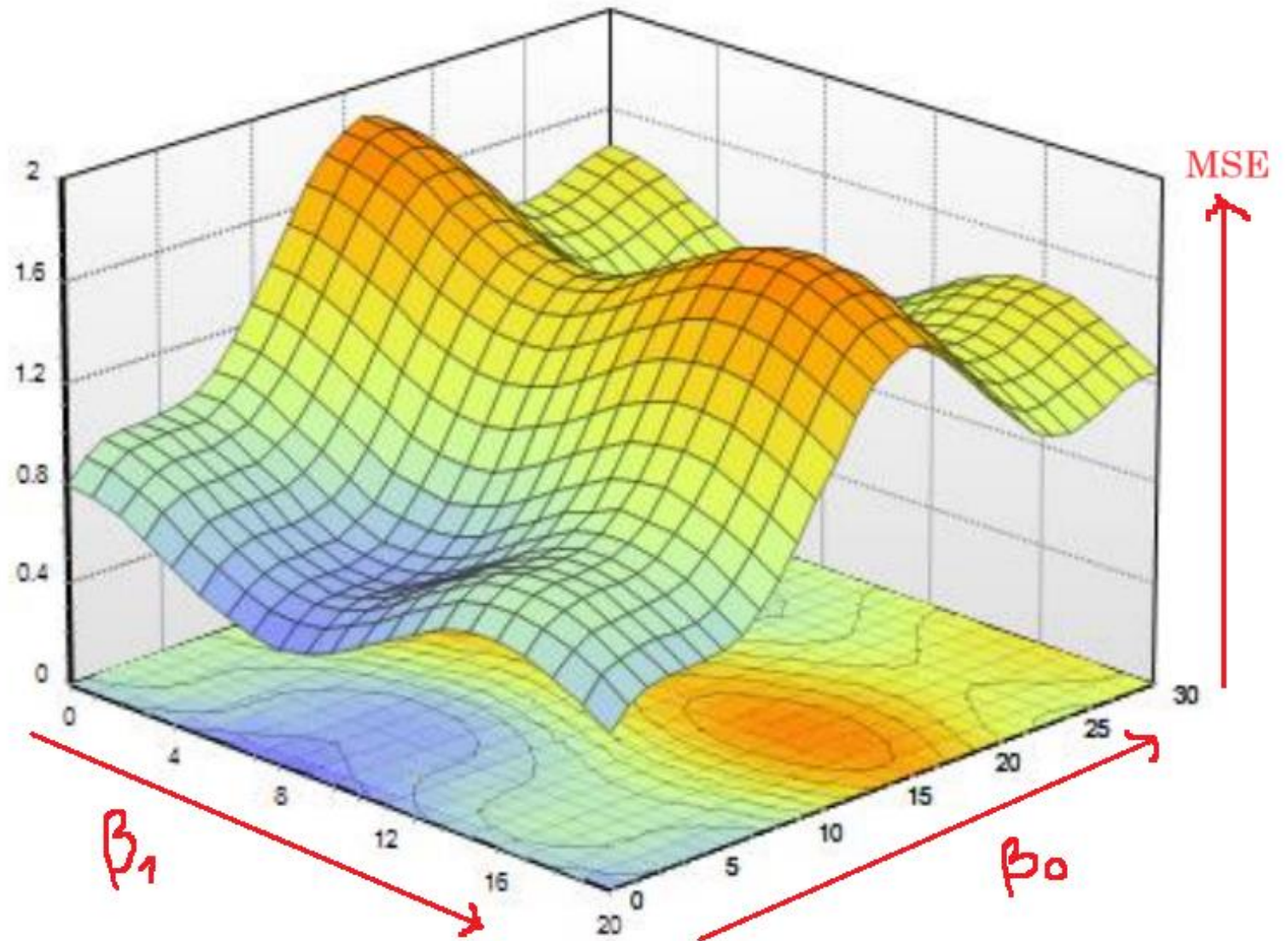
# Mean Squared Error

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$



# MSE jako funkcja

Znaleźć takie  $\beta_1, \beta_0$  żeby  
MSE było najmniejsze



Pochodne!

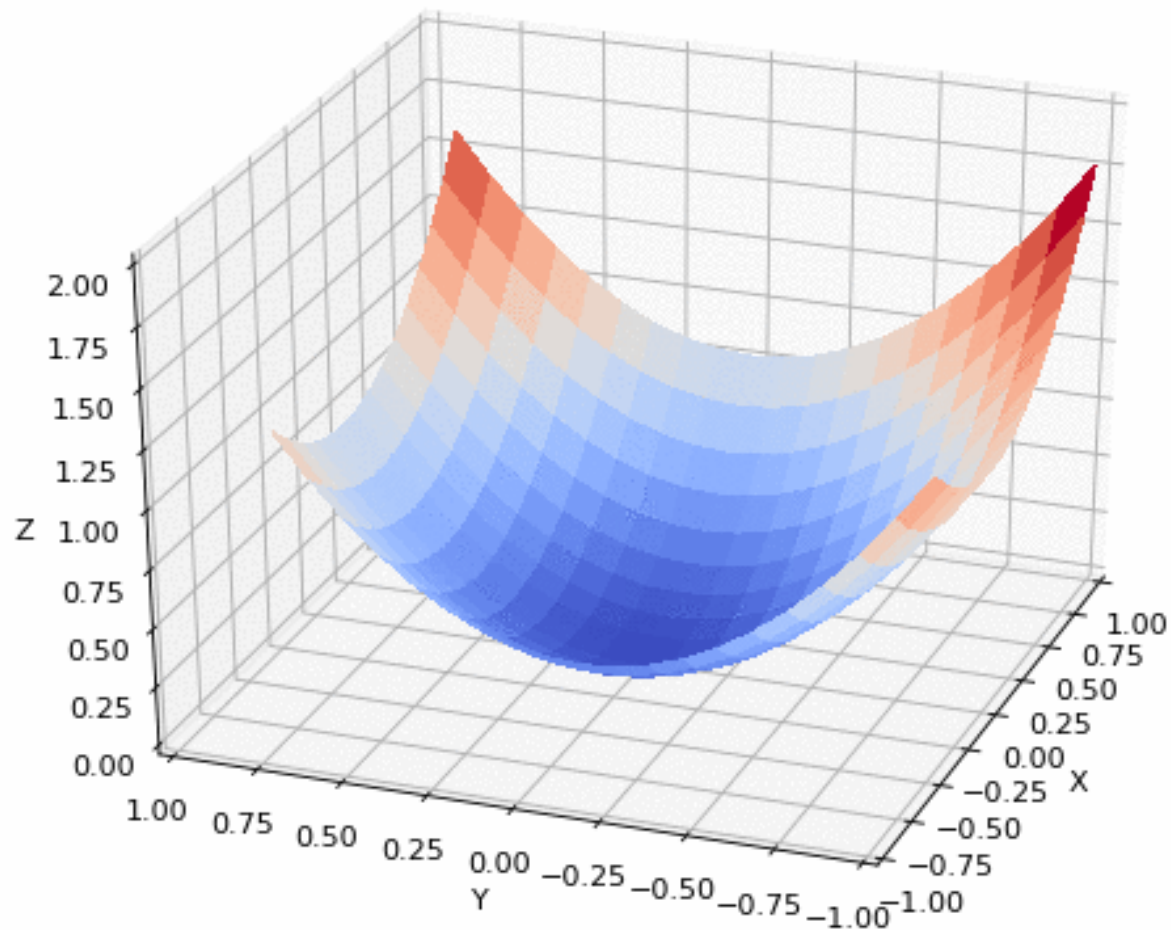
$$\frac{\partial MSE}{\partial \beta_0} = \frac{-2}{n} \sum_{i=0}^n (y_i - (\beta_0 + \beta_1 x_i))$$

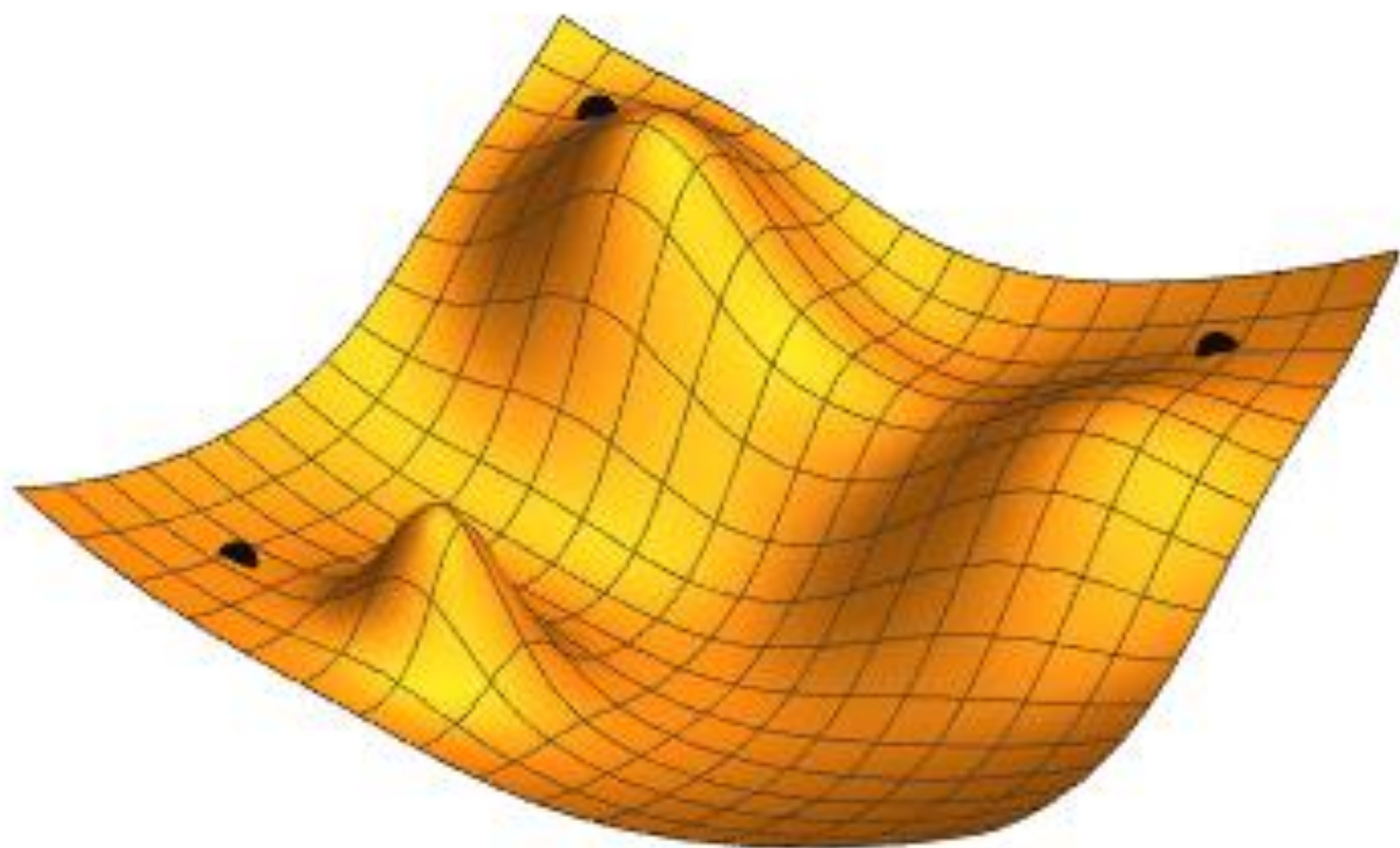
$$\frac{\partial MSE}{\partial \beta_1} = \frac{-2}{n} \sum_{i=0}^n x_i (y_i - (\beta_0 + \beta_1 x_i))$$

# Spadek gradientu (gradient descent)

Gradient = wektor wszystkich pochodnych cząstkowych

- Wskazuje nam kierunek w którym funkcja **rośnie**
- Możemy poruszać się w kierunku **przeciwnym**, aby znaleźć minimum



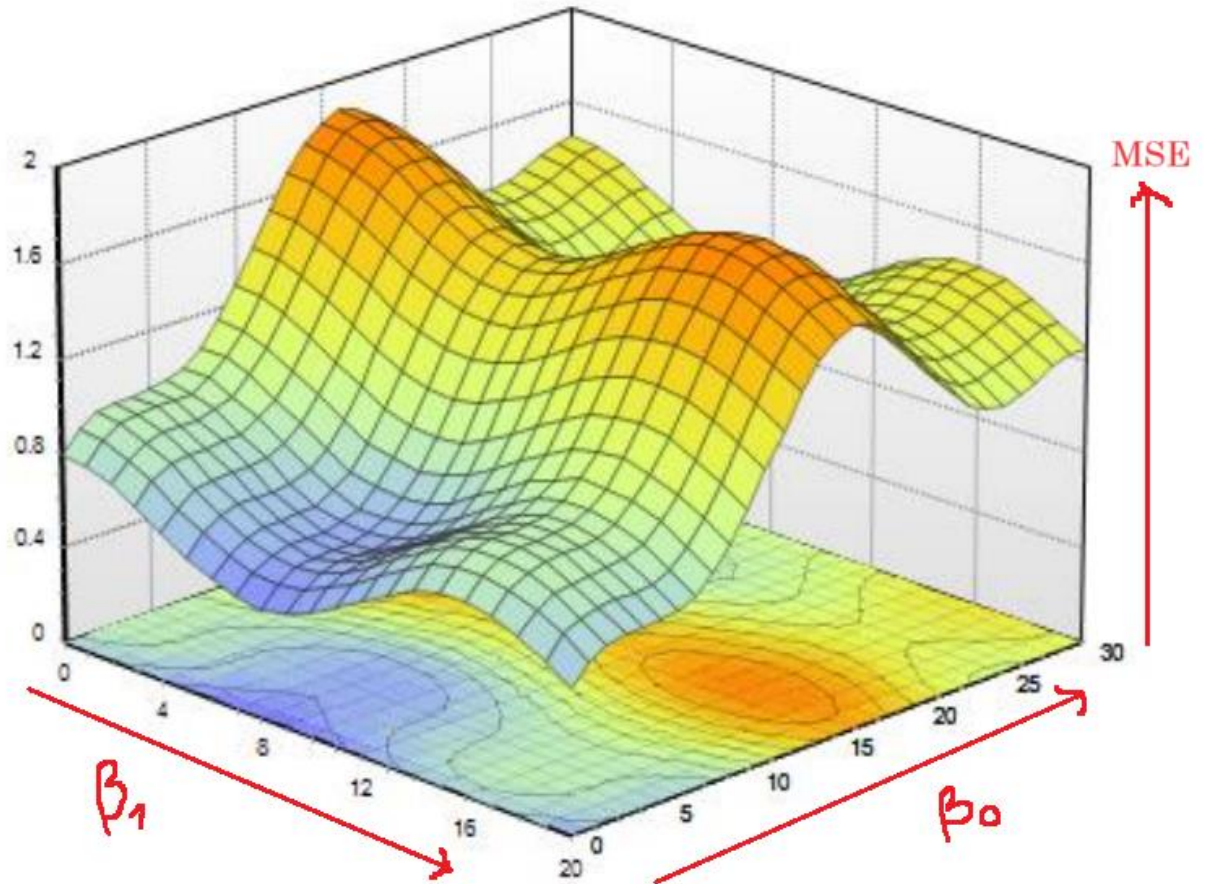




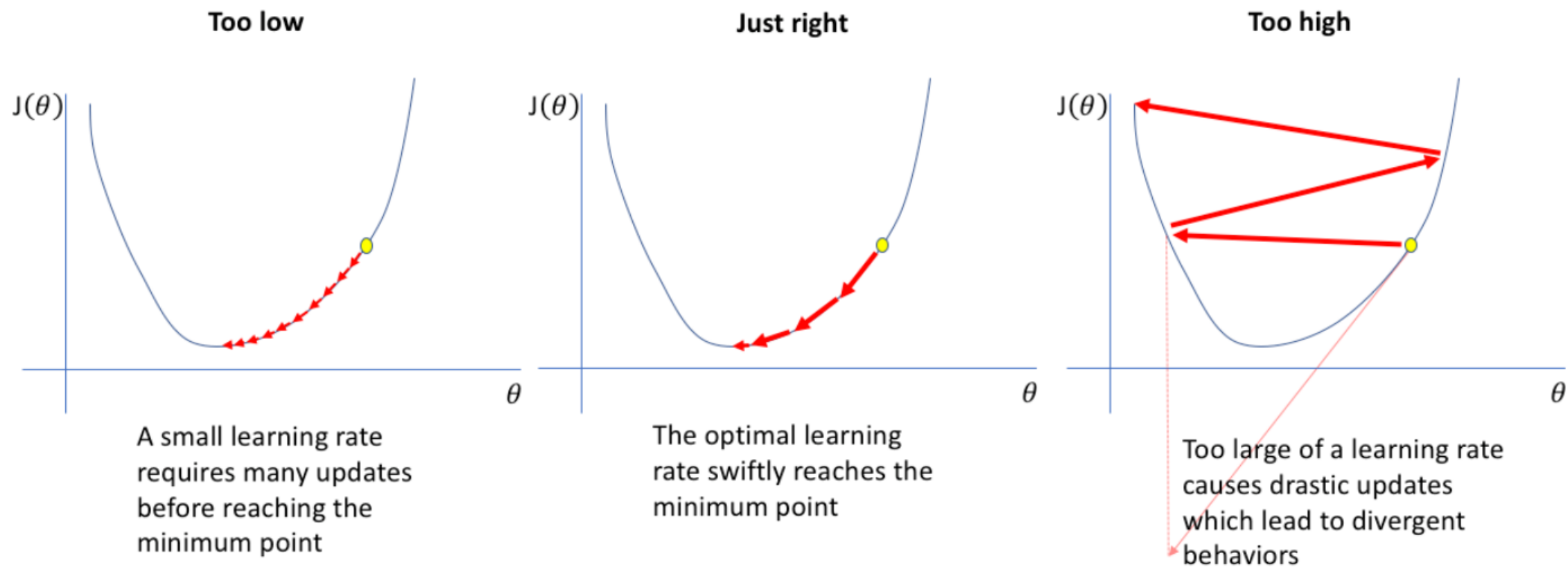
Przesuwamy się do minimum

$$\beta_0 = \beta_0 - \frac{\partial MSE}{\partial \beta_0} \cdot \alpha$$

$$\beta_1 = \beta_1 - \frac{\partial MSE}{\partial \beta_1} \cdot \alpha$$



# Współczynnik uczenia (learning rate) $\alpha$



# Więcej zmiennych

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

No	X1 transaction date	X2 house age	X3 distance to the nearest MRT station	X4 number of convenience stores	X5 latitude	X6 longitude	Y house price of unit area
1	2012.917	32	84.87882	10	24.98298	121.54024	37.9
2	2012.917	19.5	306.5947	9	24.98034	121.53951	42.2
3	2013.583	13.3	561.9845	5	24.98746	121.54391	47.3
4	2013.5	13.3	561.9845	5	24.98746	121.54391	54.8
5	2012.833	5	390.5684	5	24.97937	121.54245	43.1
6	2012.667	7.1	2175.03	3	24.96305	121.51254	32.1
7	2012.667	34.5	623.4731	7	24.97933	121.53642	40.3
8	2013.417	20.3	287.6025	6	24.98042	121.54228	46.7
9	2013.5	31.7	5512.038	1	24.95095	121.48458	18.8
10	2013.417	17.9	1783.18	3	24.96731	121.51486	22.1
11	2013.083	34.8	405.2134	1	24.97349	121.53372	41.4
12	2013.333	6.3	90.45606	9	24.97433	121.5431	58.1
13	2012.917	13	492.2313	5	24.96515	121.53737	39.3
14	2012.667	20.4	2469.645	4	24.96108	121.51046	23.8
15	2013.5	13.2	1164.838	4	24.99156	121.53406	34.3



# Algorytm spadku gradientu

1. Oblicz pochodne cząstkowe funkcji celu
2. Przemnóż pochodne przez współczynnik uczenia
3. Zaaktualizuj parametry
4. Powtarzaj dopóki wszystkie pochodne w punkcie nie będą ujemne

# Regularyzacja

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

Jeśli na przykład  $\beta_0 = 0.05$ ,  $\beta_1 = 2$ ,  $\beta_2 = 500$ , to  
wynik zależy tylko od  $x_2$

# Regularyzacja

Zmieniamy trochę funkcję straty

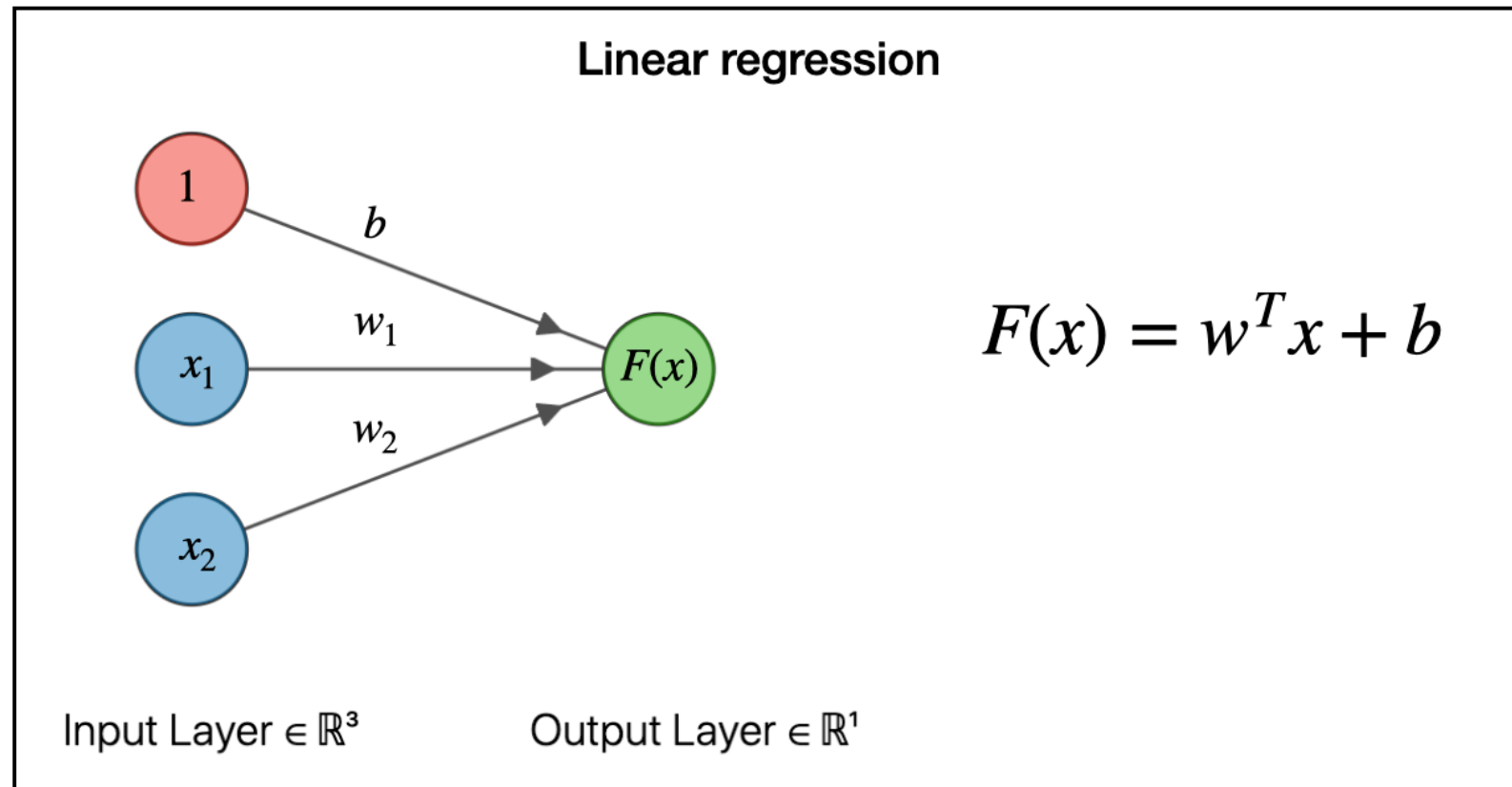
Ridge

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - pred_i)^2 + \lambda \sum_{j=0}^m \beta_j^2$$

Lasso

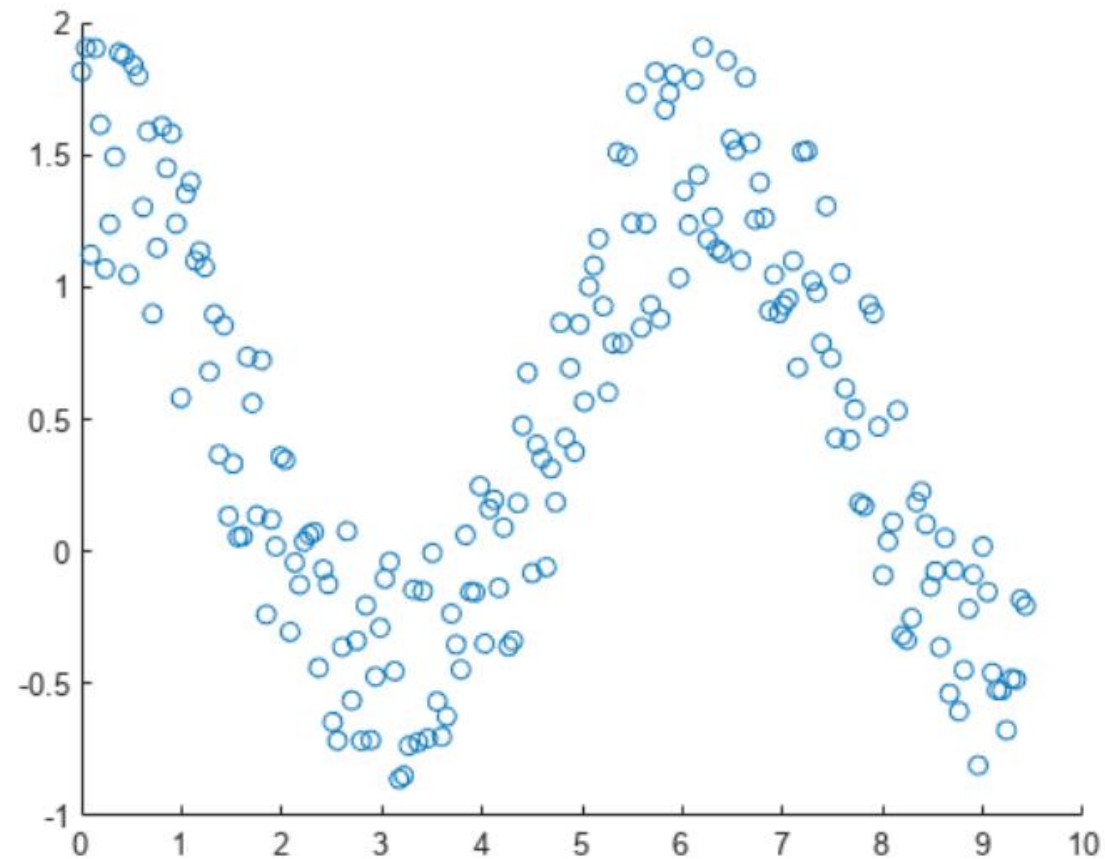
$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - pred_i)^2 + \lambda \sum_{j=0}^m |\beta_j|$$

# Po co i dlaczego?



$$F(x) = w_1 \cdot x_1 + w_2 \cdot x_2 + 1 \cdot b$$

# Inne modele regresyjne



# Regresja wielomianowa

- Bierzemy pod uwagę wszystkie możliwe iloczyny

$$y = \beta_0 + \beta_1 x_1 \longrightarrow y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \longrightarrow ???$$

# Regresja wielomianowa

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 \xrightarrow{x_2 = x_1^2} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

# Przykład – regresja kwadratowa

x1	x2	y
2	3	17
5	2	19
7	1	15



Koniec części teoretycznej



