

Giliojo mokymosi metodai

Mašininio mokymosi įvadas

2024-02-05

Vasaris

			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29			

Ką išmoksime kurso metu?

- Kas yra mašininis mokymasis
- Kas yra modeliai
- Kas yra gilieji neuroniniai tinklai
- Kaip apmokyti savo neuroninį tinklą
- Kokių tipų modeliai yra kuriami

Turinys

- Įvadas, žymėjimai, sąvokos, prog. karkasai
- Klasifikavimas, tikslo funkcija, vertinimo statistikos
- Modelių kūrimas, parametrai, hiperparametrai, reguliarizacija
- Gradientai, diferencijavimas, optimizavimo metodai, perdavimo funkcijos
- Transformacijos, įterpiniai (*embeddings*), konvoliucijos, vaizdų klasifikavimas
- Duomenų augmentacijos, žinių perkėlimas (*transfer learning*), aktyvacijų žemėlapiai
- Lokalizacija ir segmentacija
- Autoenkoderiai, vaizdų generavimas
- Transformeriai, dėmesio sutelkimas,
- Natūralios kalbos apdorojimas
- Generatyviniai priešiški tinklai, difuzijos modeliai

Ką būtų naudinga žinoti?

Programavimas (*frameworks*):

- Python (dirbsime su NumPy, PyTorch, TensorFlow, JAX)

Matematika:

- Algebra (tiesinės lygtys, vektoriai, matricos)
- Matematinė analizė (išvestinės)
- Optimizavimo metodai (gradientinis nusileidimas)

Kursai:

- Skaitmeninis intelektas ir sprendimų priėmimas (prof. dr. Olga Kurasova)
- Optimizavimo metodai

Mašininio mokymosi įvadas

Mašininis mokymasis – dirbtinio intelekto studijų sritis, susijusi su statistinių algoritmų, kurie sugeba efektyviai apibendrinti (*generalize*) ir atlikti užduotis be aiškių nurodymų (*without explicit instructions*), kūrimu (Koza et al. 1996).

Koza, J.R., Bennett, F.H., Andre, D., Keane, M.A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In: Gero, J.S., Sudweeks, F. (eds) *Artificial Intelligence in Design '96*. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0279-4_9

Mašininio mokymosi įvadas

Mašininis mokymasis – dirbtinio intelekto studijų sritis, susijusi su statistinių algoritmų, kurie sugeba efektyviai apibendrinti (*generalize*) ir atlikti užduotis be aiškių nurodymų (*without explicit instructions*), kūrimu (Koza et al. 1996).



Koza, J.R., Bennett, F.H., Andre, D., Keane, M.A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In: Gero, J.S., Sudweeks, F. (eds) *Artificial Intelligence in Design '96*. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0279-4_9

Mašininio mokymosi įvadas

Mašininis mokymasis – dirbtinio intelekto studijų sritis, susijusi su statistinių algoritmų, kurie sugeba efektyviai apibendrinti (*generalize*) ir atlikti užduotis be aiškių nurodymų (*without explicit instructions*), kūrimu (Koza et al. 1996).



```
def has_ears(image):  
    # find oval in image center  
    circle = image.matchOval(image, x=100, y=100, w=30, h=20)  
  
    # find two triangles  
    triangle1 = image.matchTriangle(circle.x-20, circle.y)  
    triangle2 = image.matchTriangle(circle.x+20, circle.y)  
  
    if triangle1 and triangle2:  
        return True  
    else:  
        return False
```

Koza, J.R., Bennett, F.H., Andre, D., Keane, M.A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In: Gero, J.S., Sudweeks, F. (eds) *Artificial Intelligence in Design '96*. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0279-4_9

Kokias problemas gali spręsti giliojo mokymosi metodai?

- Daug pavyzdžių (*labeled examples*)
- Savybės ir požymiai (*patterns*) duomenyse
- Sunkiai išreiškiami sąryšiai tarp kintamųjų (*complex relationships*)

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Modelis

Modelis – realaus proceso/sistemas supaprastinimas.

x	y
2	19.6
4.5	44.1
10	98



Modelis

Modelis – realaus proceso/sistemas supaprastinimas.

x	y
2	19.6
4.5	44.1
10	98

$$y = ?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98

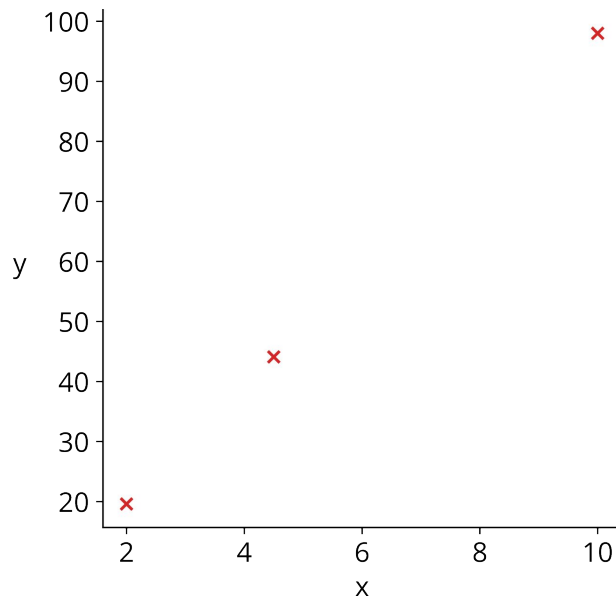
$$y = ?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



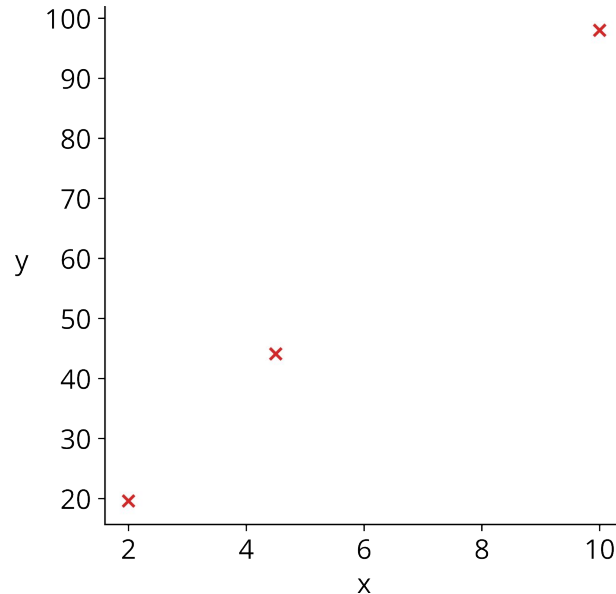
$y = ?$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



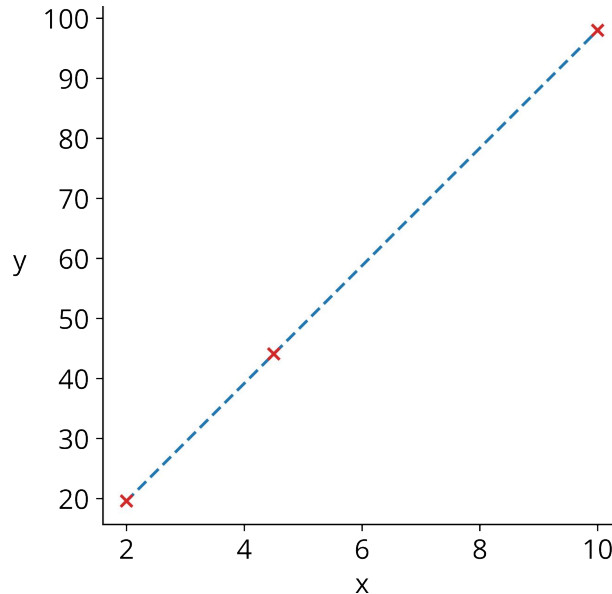
$$y = ax$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

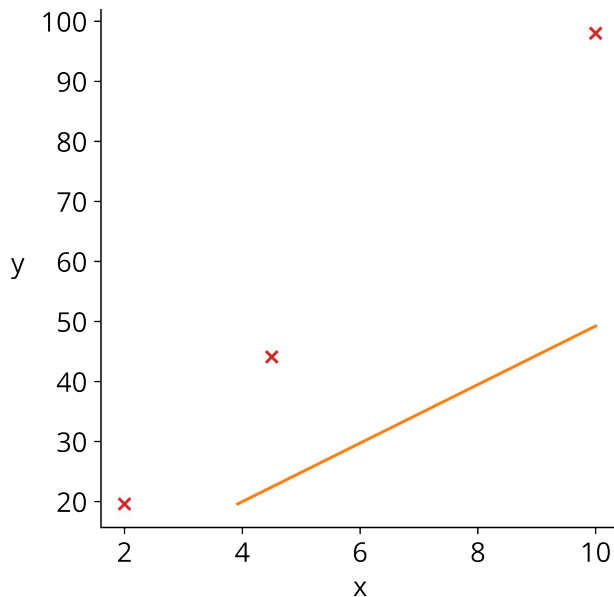
$$a = ?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

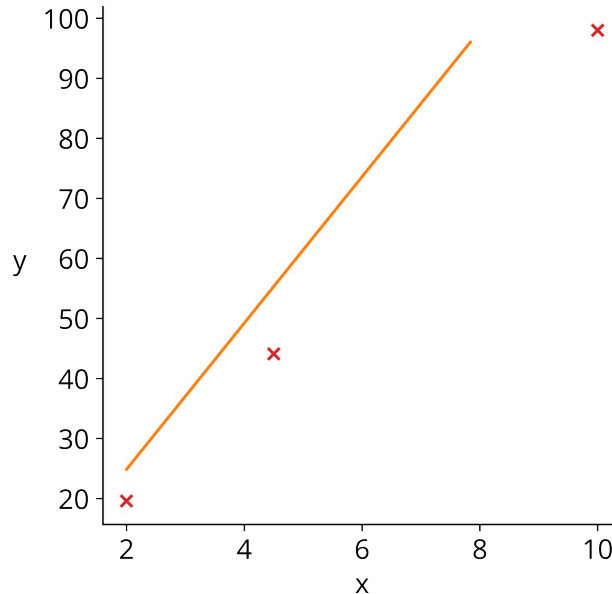
$$a = 5?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

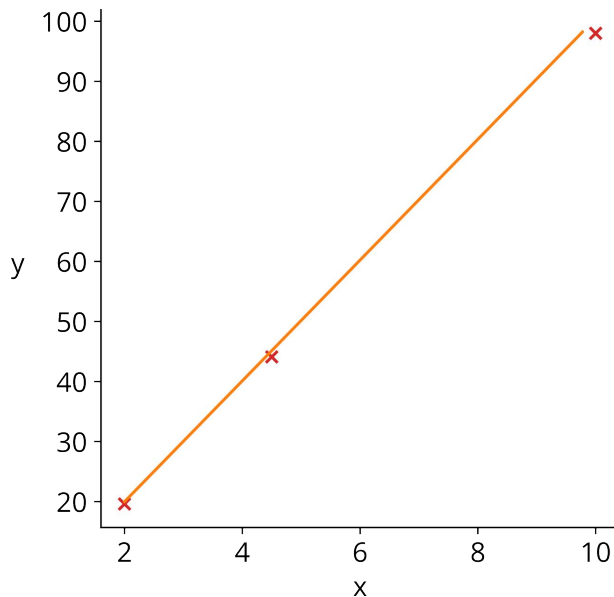
$$a = 12?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

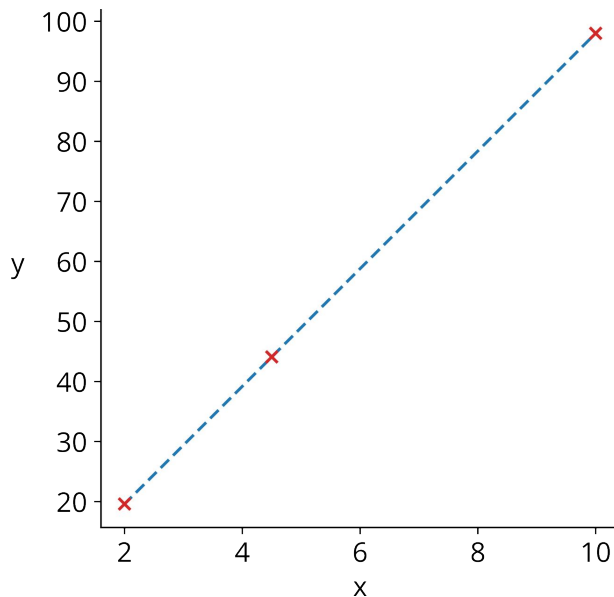
$$a = 10?$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

$$a = 9.8$$

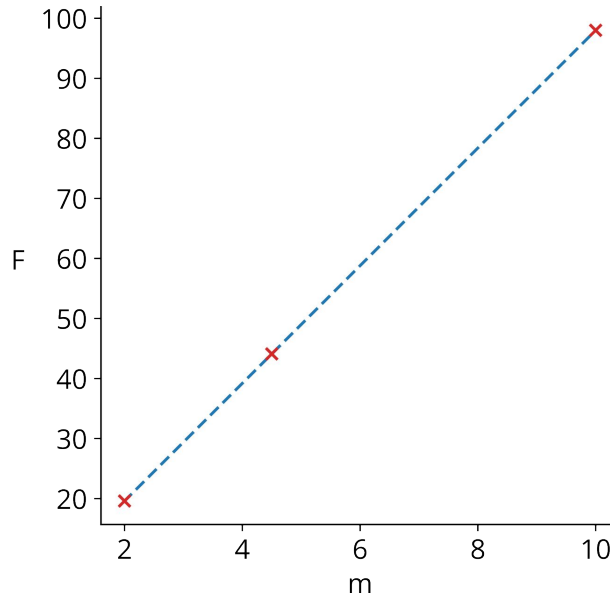
$$y = 9.8x$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

m	F
2	19.6
4.5	44.1
10	98



$$F = mg$$

$$g = 9.8$$

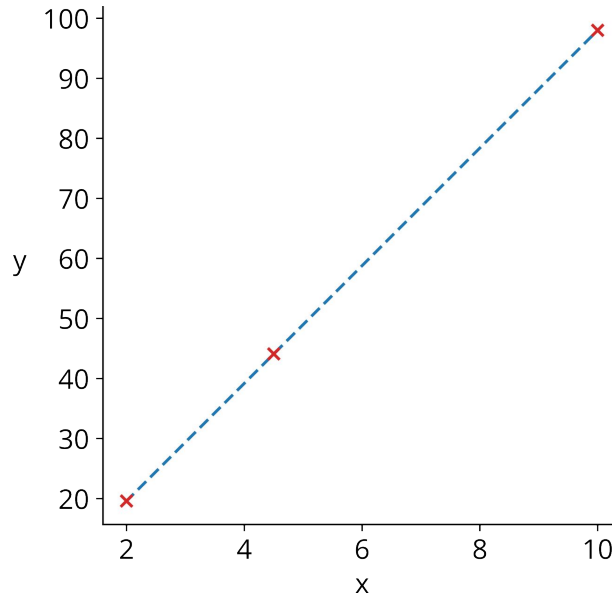
$$F = 9.8 m$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$y = ax$$

$$a = 9.8$$

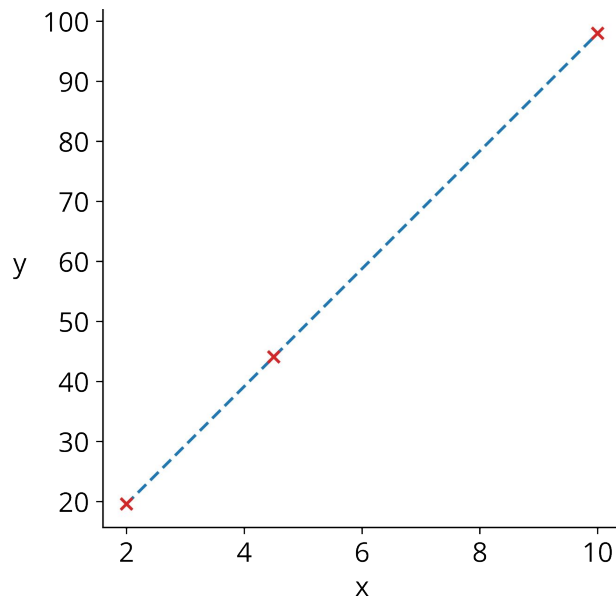
$$y = 9.8x$$

Modelis

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98




$$\hat{y} = ax$$

$$a = 9.8$$

$$\hat{y} = 9.8x$$

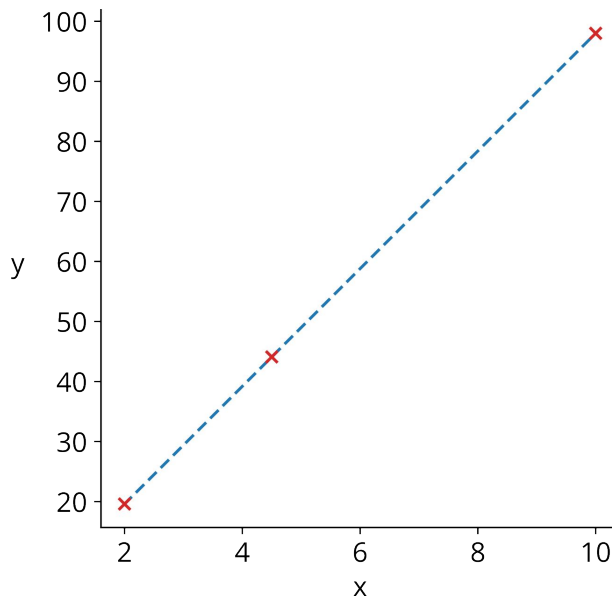
Modelis

(x, y) – duomenys 
 \hat{y} – spėjimas (*prediction*)

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y
2	19.6
4.5	44.1
10	98



$$\hat{y} = ax$$

$$a = 9.8$$

$$\hat{y} = 9.8x$$

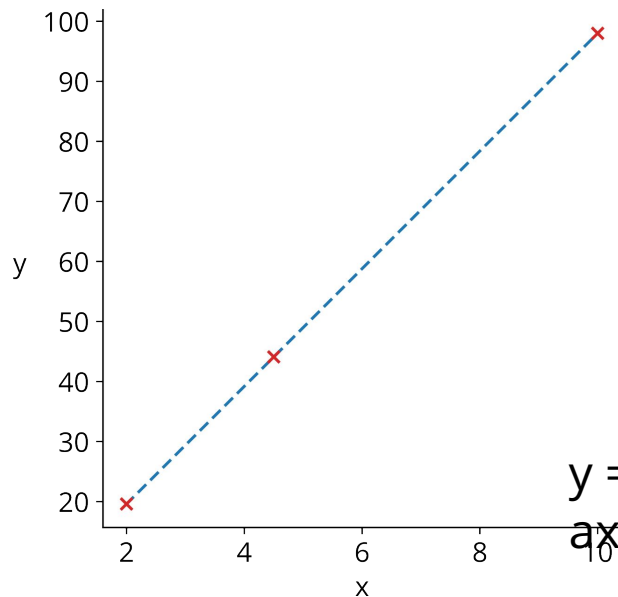
Modelis

(x, y) – duomenys
 \hat{y} – spėjimas (*prediction*)

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai norime tiksliai žinoti kaip vyksta *duomenis generuojantis procesas*.

x	y	z
2	19.6	?
4.5	44.1	?
10	98	?



$$\hat{y} = ax$$

$$a = 9.8$$

$$\hat{y} = 9.8x$$

$$y = (1 + 1/(1 - z^2) z z^T / b^2) (1 - z^2)^{-1/2}$$

$$b \approx 2.99 \times 10^8$$



Modelis

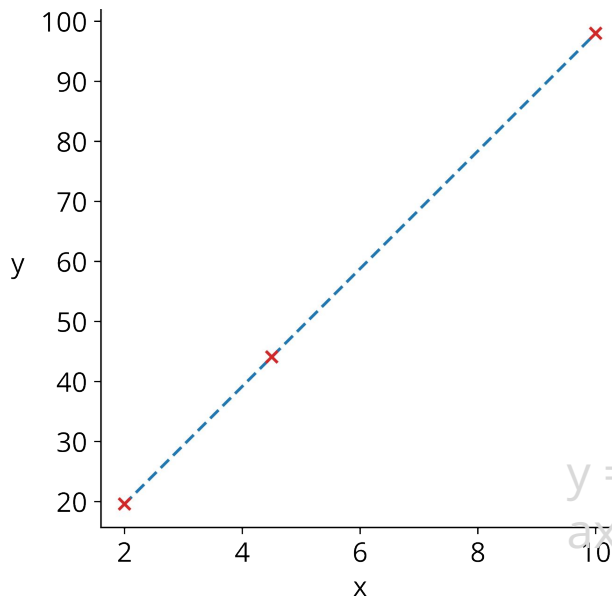
(x, y) – duomenys
 \hat{y} – spėjimas (*prediction*)

Modelis – realaus proceso/sistemos supaprastinimas.

Dažnai *užtenka aproksimuoti* kaip vyksta *duomenis generuojantis procesas*.

x	y	z
2	19.6	?
4.5	44.1	?
10	98	?

$$\hat{y} \approx y$$



$$\hat{y} = ax$$

$$a = 9.8$$

$$\hat{y} = 9.8x$$

$$y = (1 + 1/(1 - z^2) z z^T / b^2) (1 - z^2)^{-1/2}$$

$$b \approx 2.99 \times 10^8$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?



x	y
2	19.6
4.5	44.1
10	98

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$y_i - \hat{y}(x_i)$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\left| y_i - \hat{y}(x_i) \right|$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\left(y_i - \hat{y}(x_i)\right)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\sum_i \left(y_i - \hat{y}(x_i) \right)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{Jei } \mathcal{L} = 0 \quad \text{tai } y_i - \hat{y}(x_i) = 0 \text{ kiekvienam } i$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{Jei } \mathcal{L} = 0 \quad \text{tai } y_i = \hat{y}(x_i) \quad \text{kiekvienam } i$$

Nuostolių funkcija

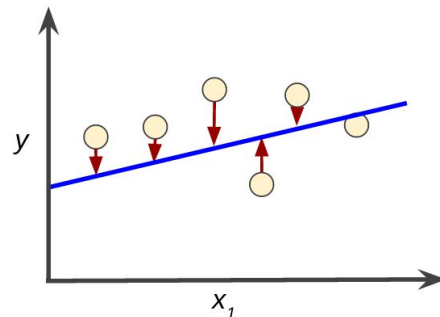
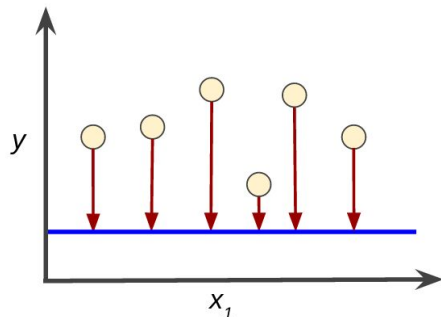
Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{Jei } \mathcal{L} = 0 \text{ tai } y_i = \hat{y}(x_i) \text{ kiekvienam } i$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 550 595 620"/>$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = a x$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = a x \quad \mathcal{L} = (y_1 - a \cdot x_1)^2 + (y_2 - a \cdot x_2)^2 + (y_2 - a \cdot x_2)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = a x \quad \mathcal{L} = (19.6 - a \cdot 2)^2 + (44.1 - a \cdot 4.5)^2 + (98 - a \cdot 10)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2$$

Nuostolių funkcija

Kas yra gera aproksimacija?

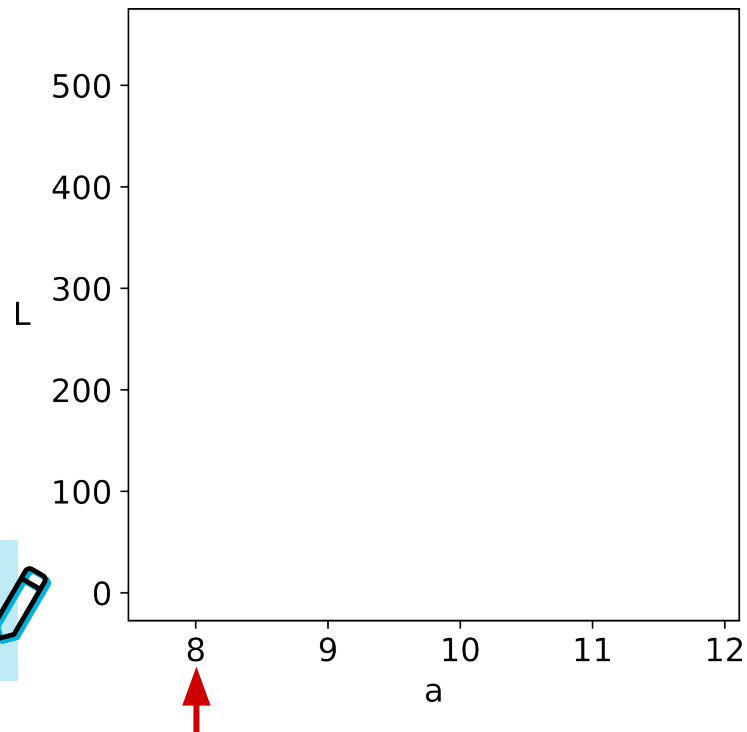
Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2$$



Nuostolių funkcija

Kas yra gera aproksimacija?

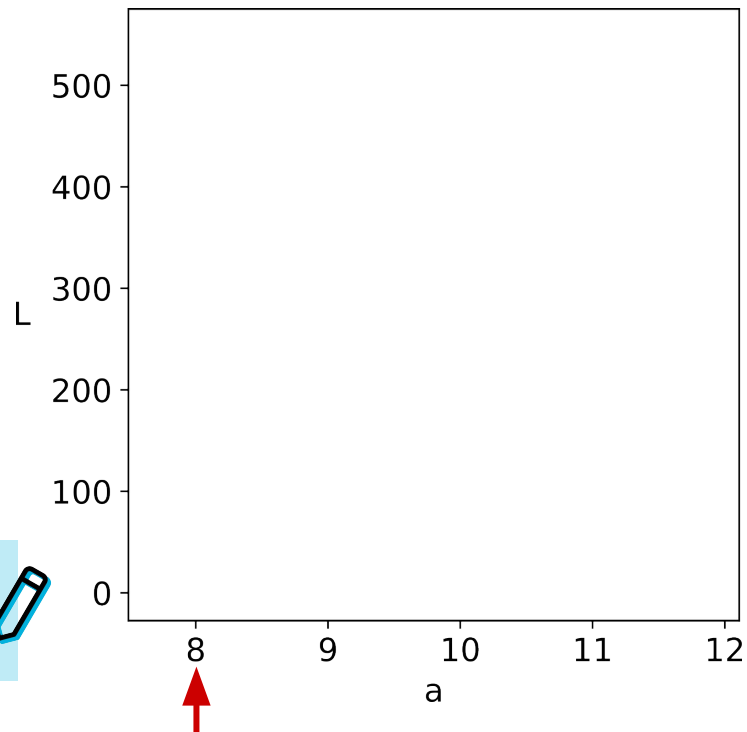
Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss}$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$



Nuostolių funkcija

Kas yra gera aproksimacija?

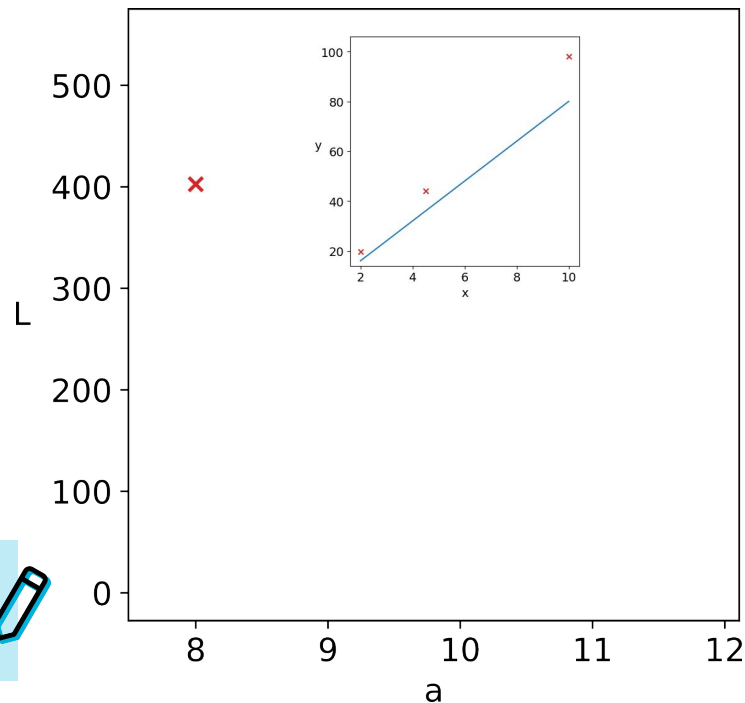
Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

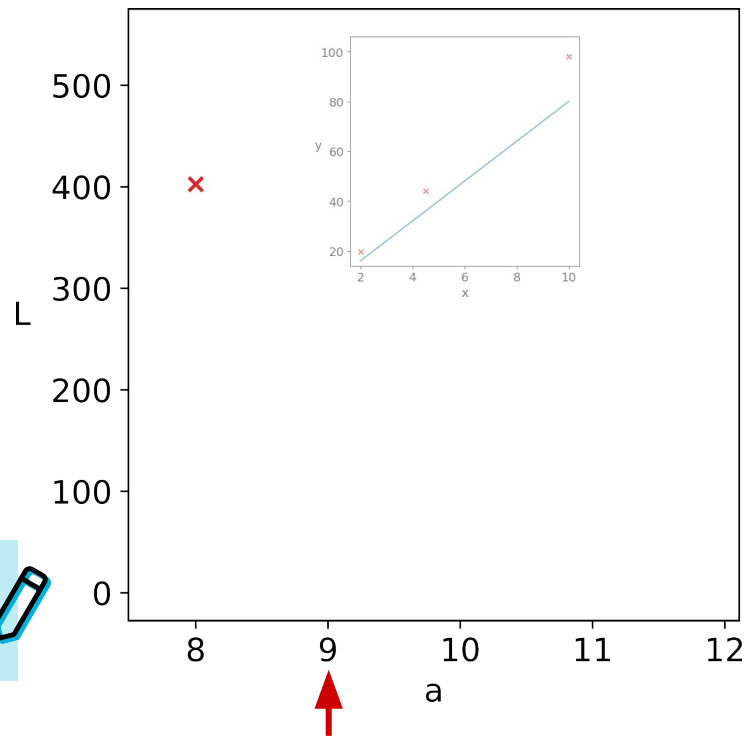
Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

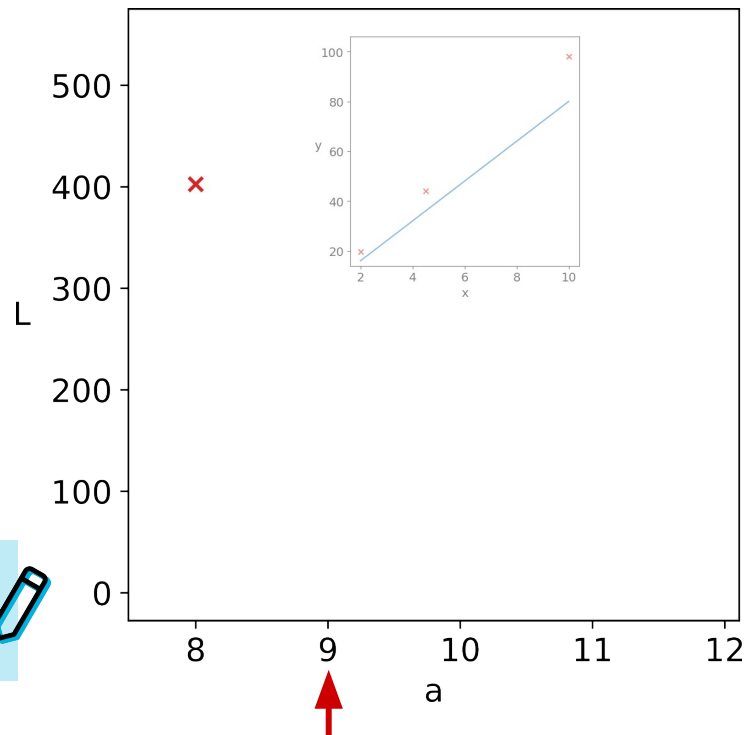
Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

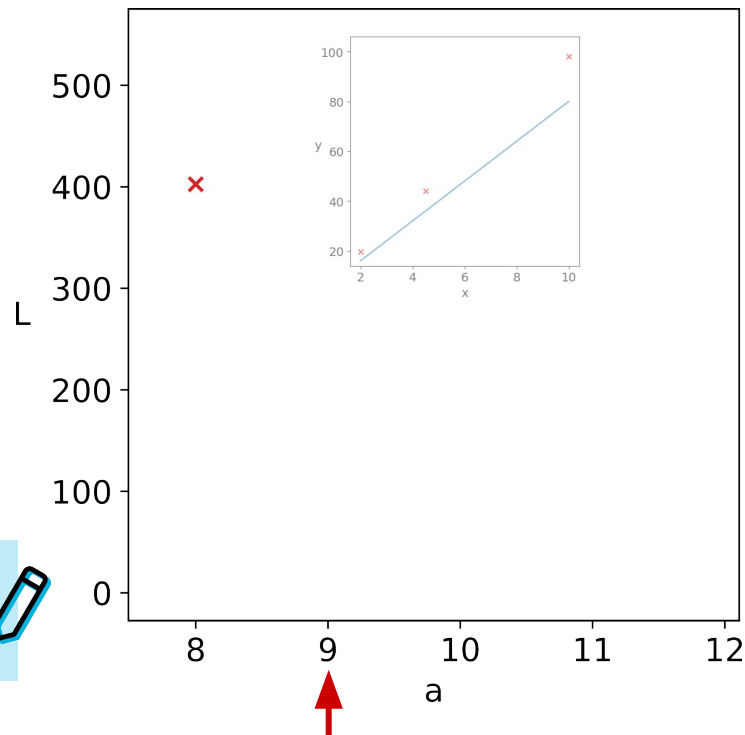
Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

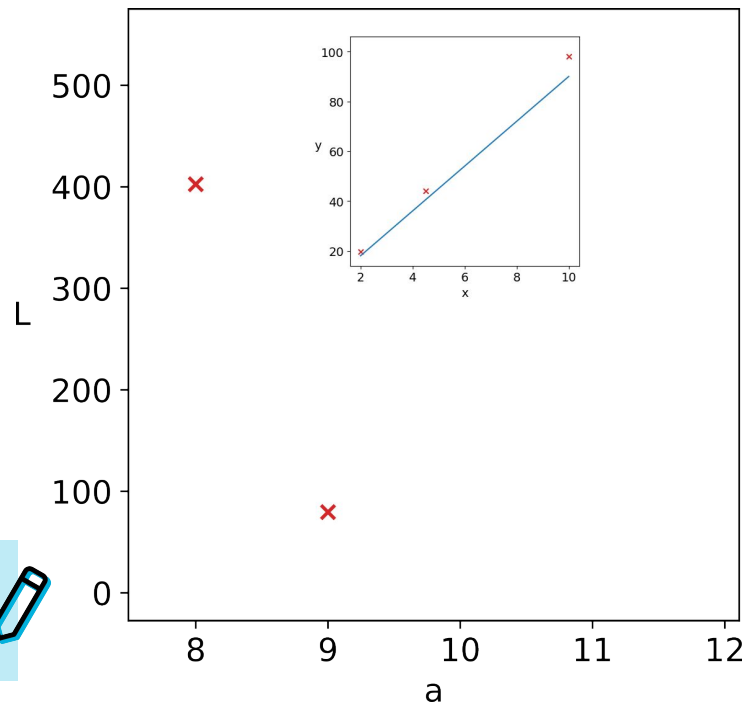
Atstumas tarp mūsų modelio ir duomenų:

x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

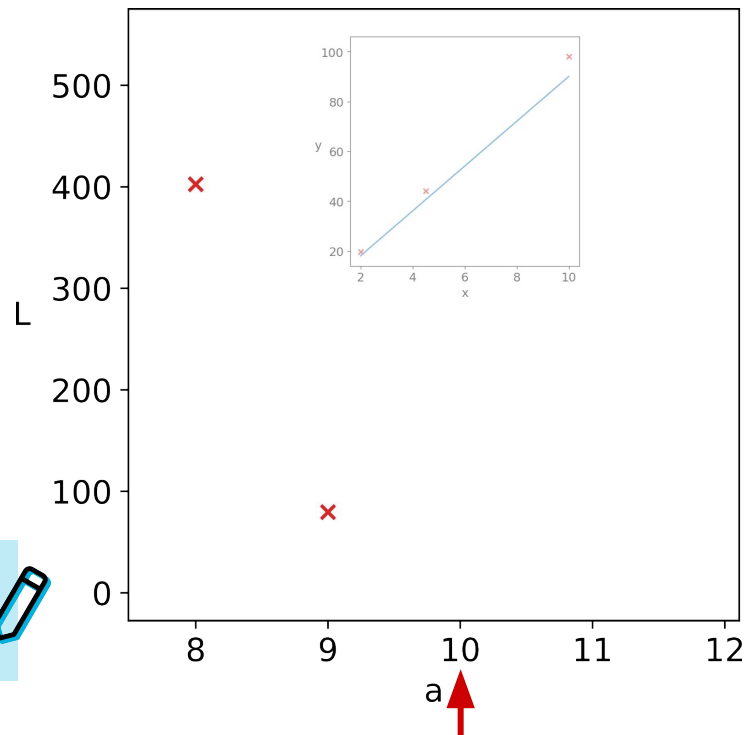
x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i (y_i - \hat{y}(x_i))^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$

$$\hat{y} = 10x$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

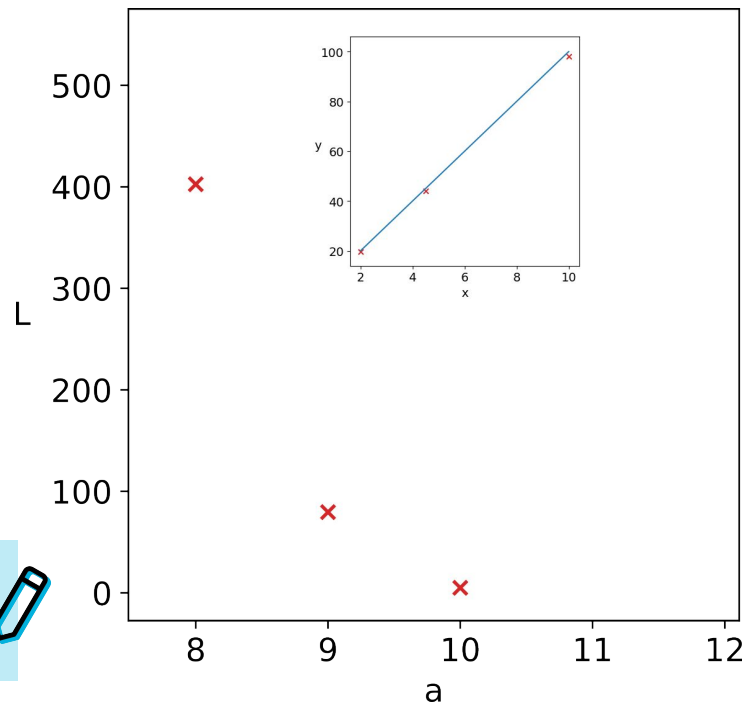
x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 622"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$

$$\hat{y} = 10x \quad \mathcal{L} = (19.6 - 10 \cdot 2)^2 + (44.1 - 10 \cdot 4.5)^2 + (98 - 10 \cdot 10)^2 = 4.97$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

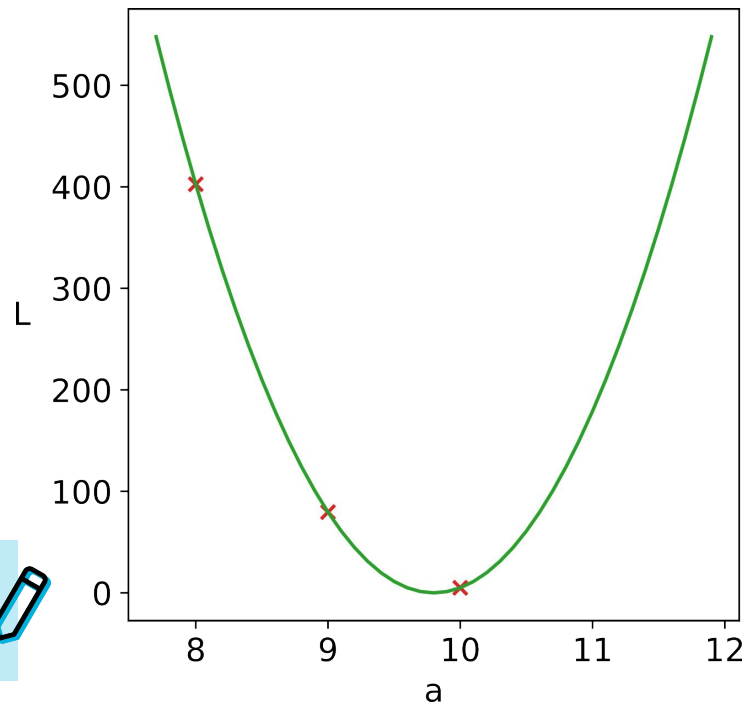
x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss} \quad \img alt="pencil icon" data-bbox="565 548 595 620"/>$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$

$$\hat{y} = 10x \quad \mathcal{L} = (19.6 - 10 \cdot 2)^2 + (44.1 - 10 \cdot 4.5)^2 + (98 - 10 \cdot 10)^2 = 4.97$$



Nuostolių funkcija

Kas yra gera aproksimacija?

Ar galima skaitiškai įvertinti „gerumą“?

Atstumas tarp mūsų modelio ir duomenų:

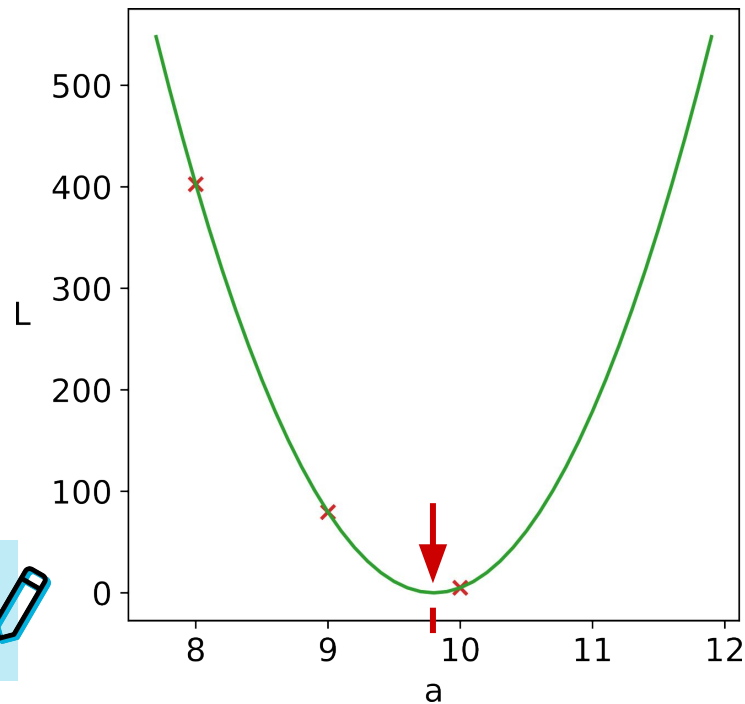
x	y
2	19.6
4.5	44.1
10	98

$$\mathcal{L} = \sum_i \left(y_i - \hat{y}(x_i) \right)^2 \quad \text{L2 Loss}$$

$$\hat{y} = 8x \quad \mathcal{L} = (19.6 - 8 \cdot 2)^2 + (44.1 - 8 \cdot 4.5)^2 + (98 - 8 \cdot 10)^2 = 402.57$$

$$\hat{y} = 9x \quad \mathcal{L} = (19.6 - 9 \cdot 2)^2 + (44.1 - 9 \cdot 4.5)^2 + (98 - 9 \cdot 10)^2 = 79.52$$

$$\hat{y} = 10x \quad \mathcal{L} = (19.6 - 10 \cdot 2)^2 + (44.1 - 10 \cdot 4.5)^2 + (98 - 10 \cdot 10)^2 = 4.97$$



Gradientinis nusileidimas

Kaip rasti tinkamą parametro **a** reikšmę?

$$\hat{y} = ax$$



Gradientinis nusileidimas

Kaip rasti minimumą?

\mathcal{L}

Gradientinis nusileidimas

Kaip rasti minimumą? $\min \mathcal{L}$

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$ $a^* = \arg \min_a \mathcal{L}(a)$

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$



Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

\mathcal{L}'

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

\mathcal{L}' $\mathcal{L}'_x?$ $\mathcal{L}'_a?$

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

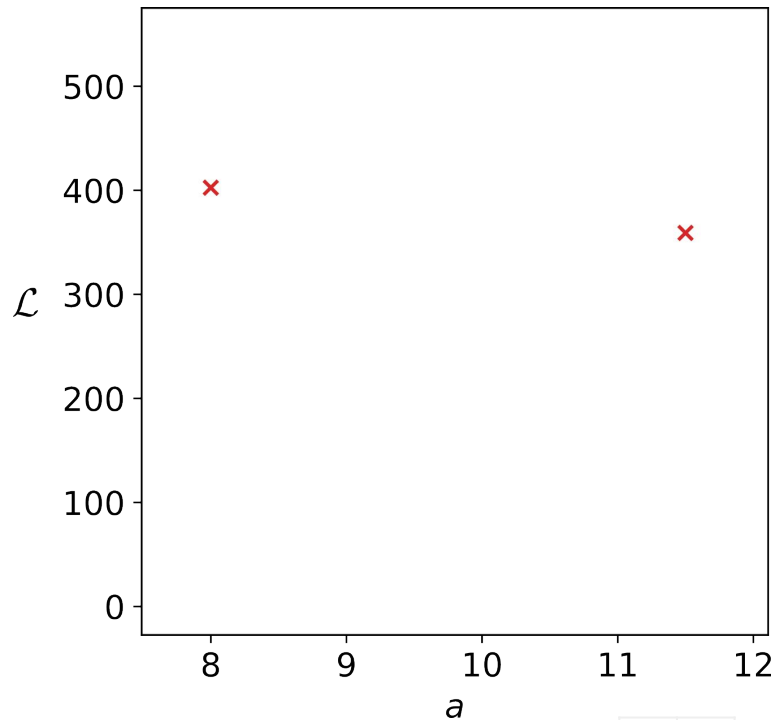
$$\frac{\partial \mathcal{L}}{\partial a}$$

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a}$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

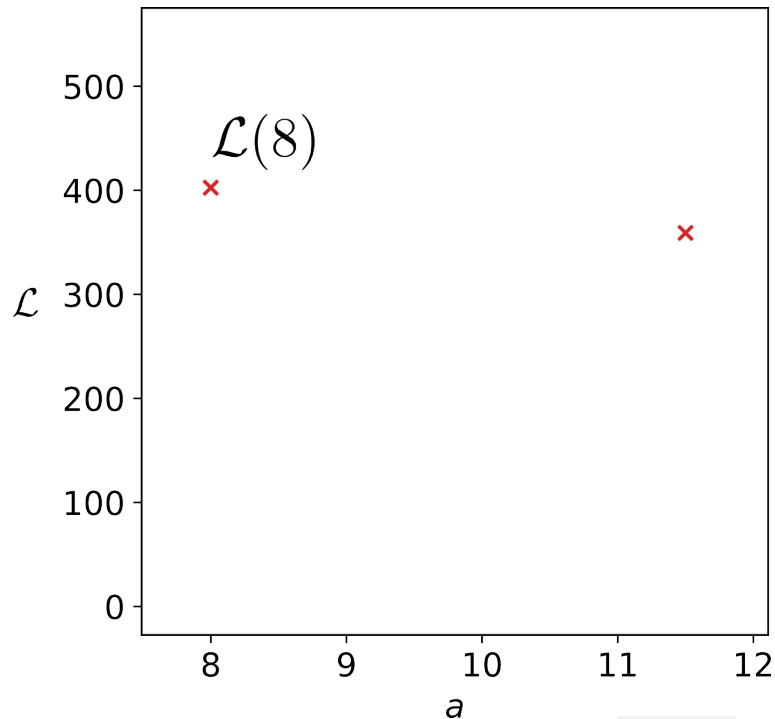
Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\mathcal{L}(8) \approx 402$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

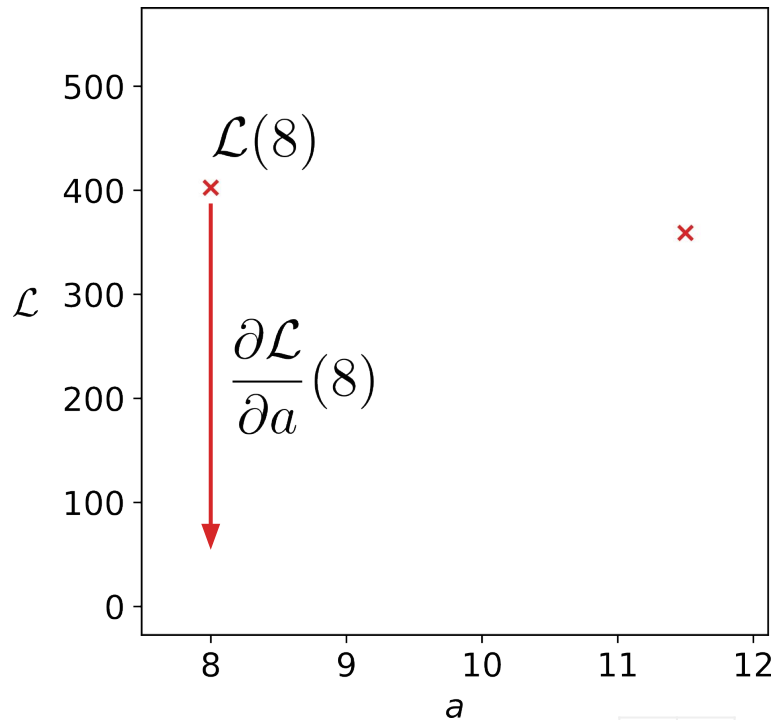
Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

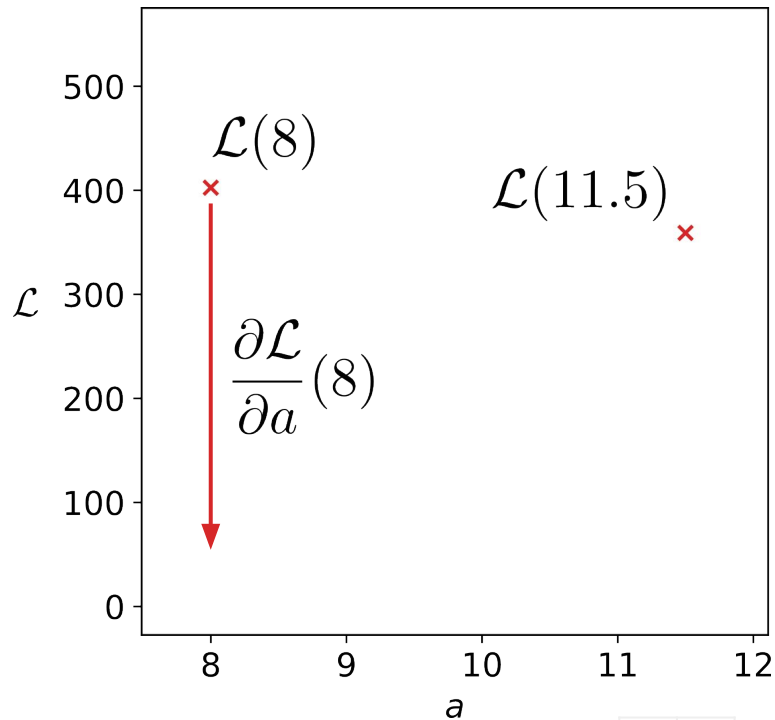
Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

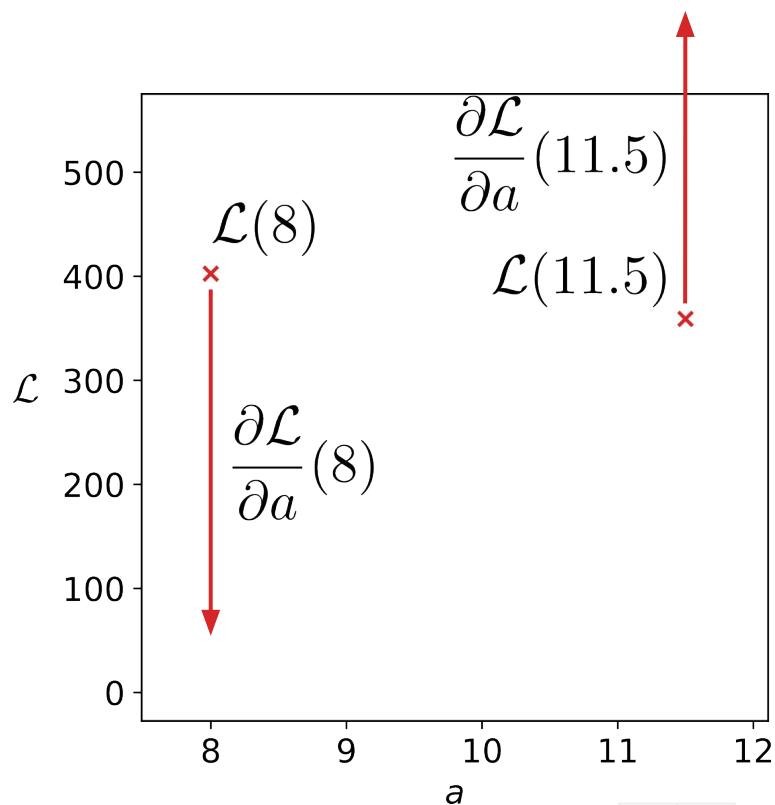
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(11.5) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

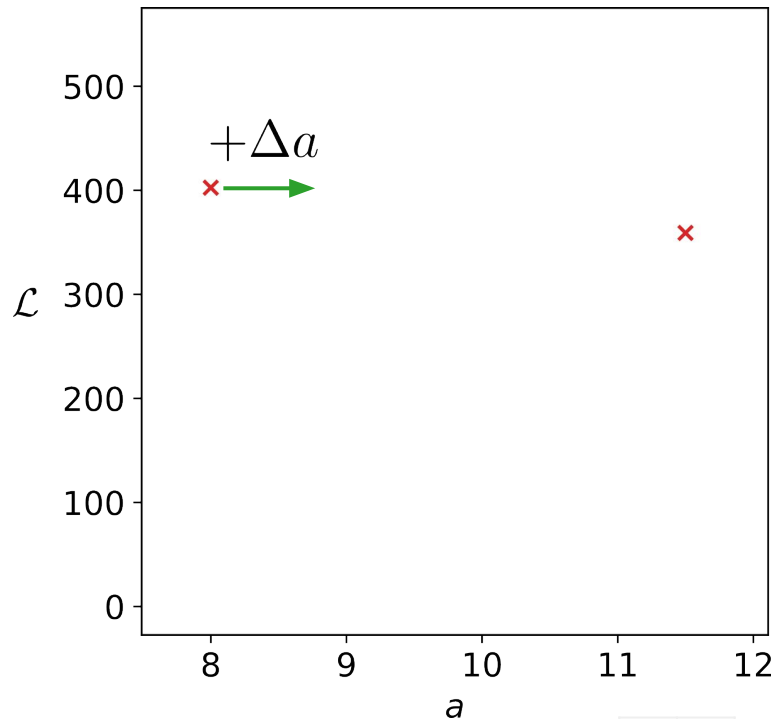
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

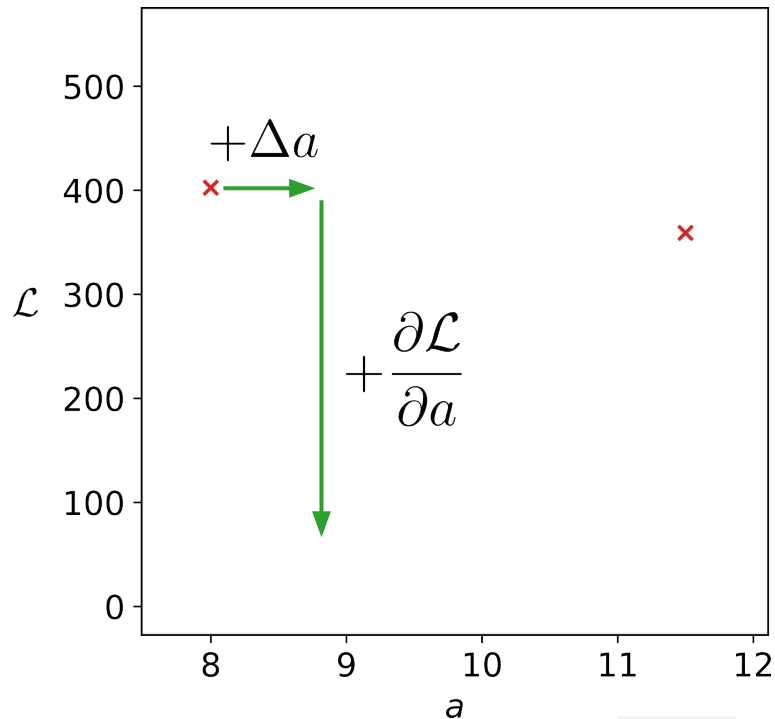
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

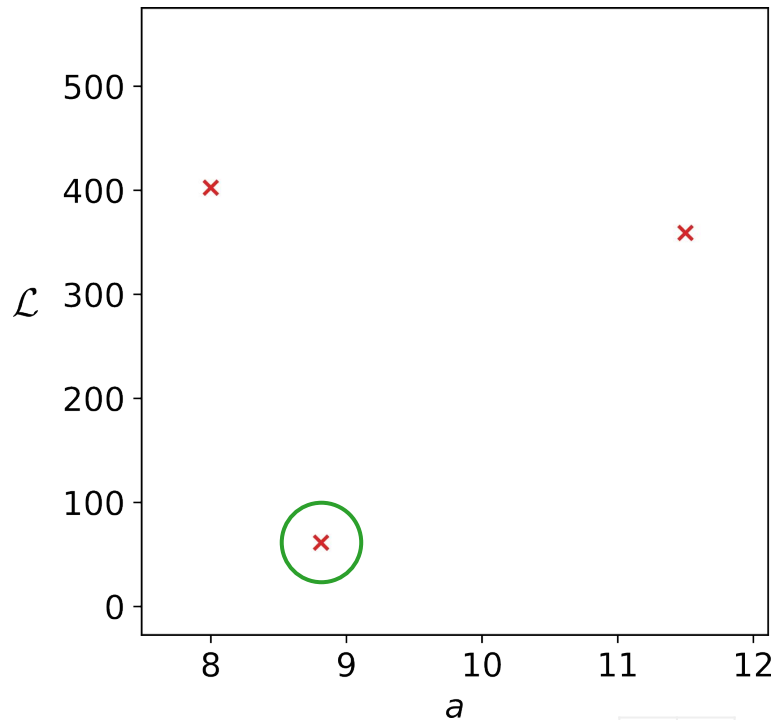
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

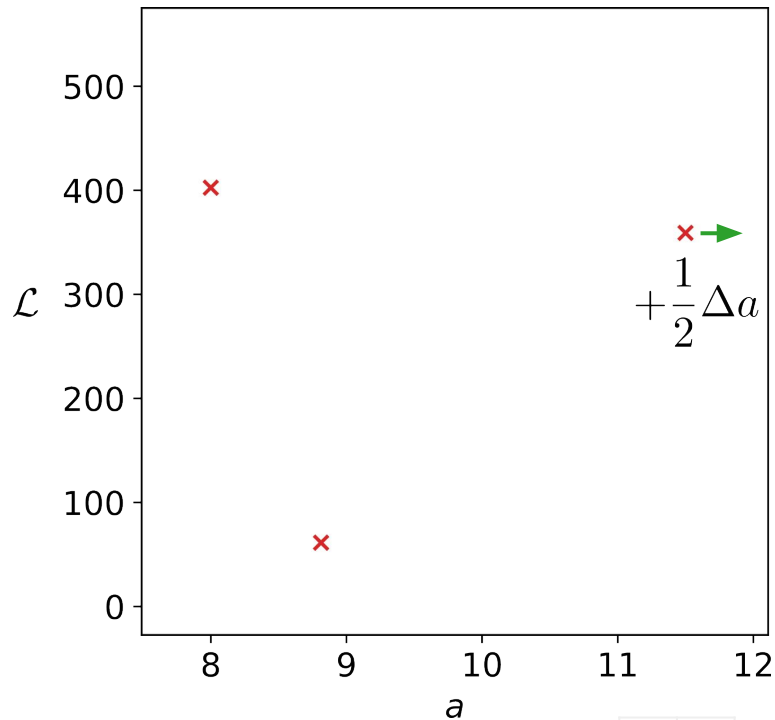
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

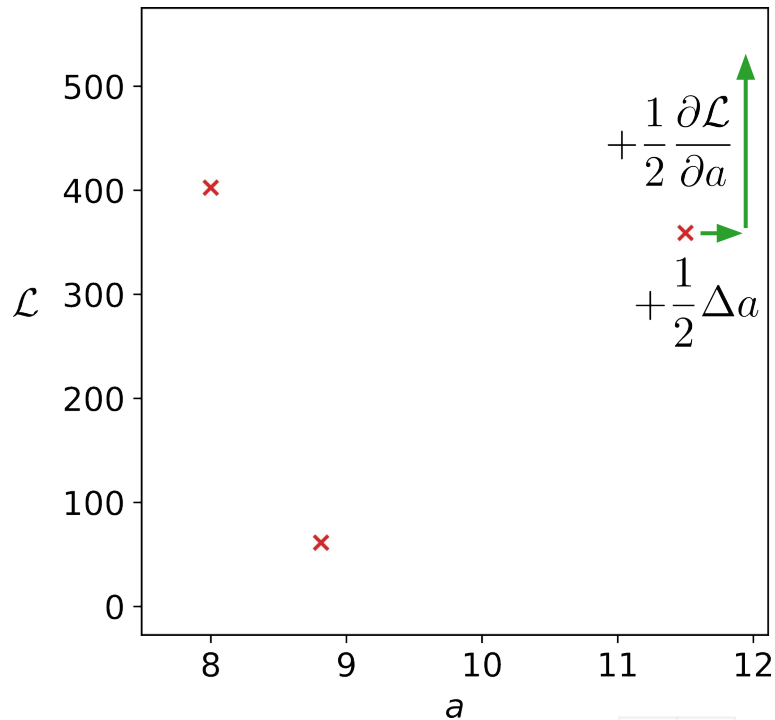
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

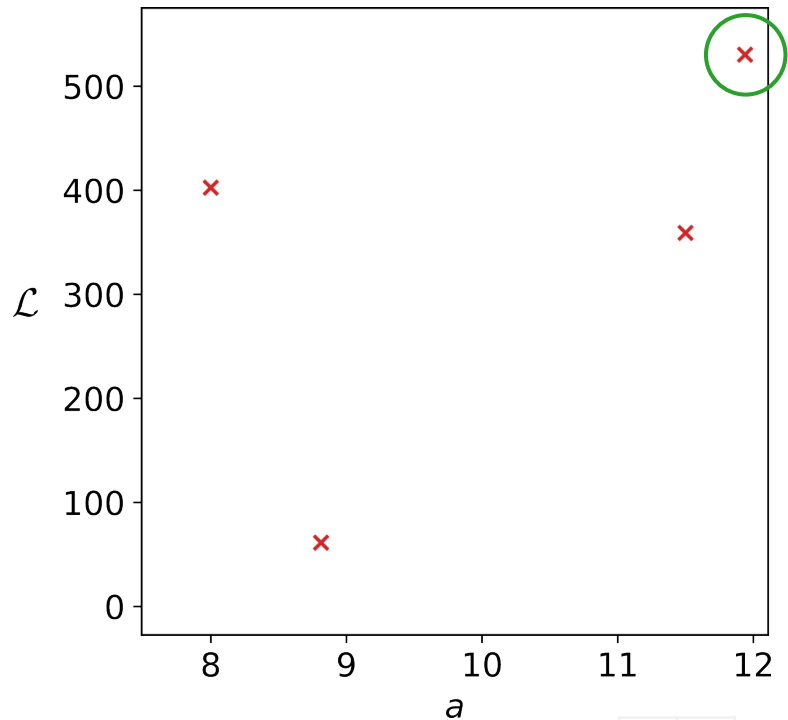
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

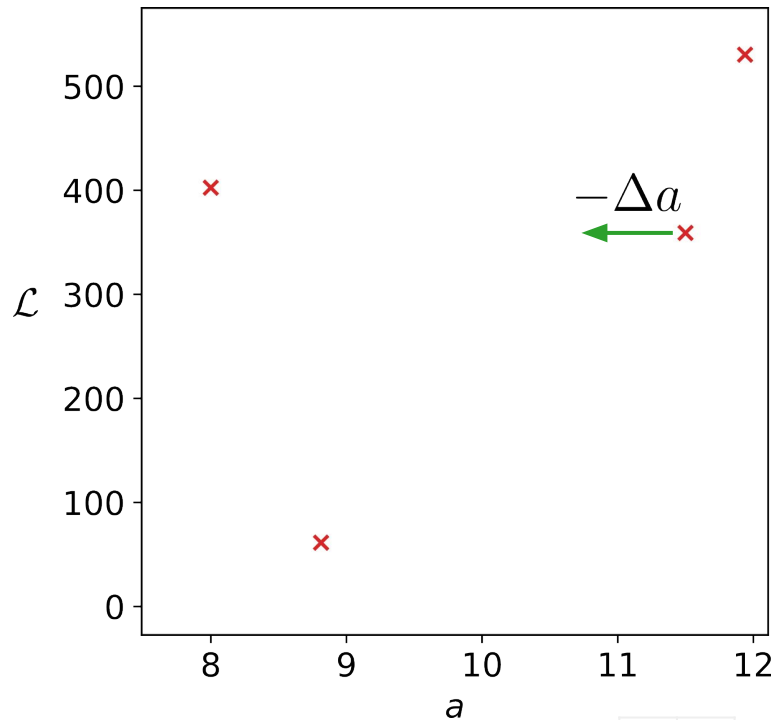
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

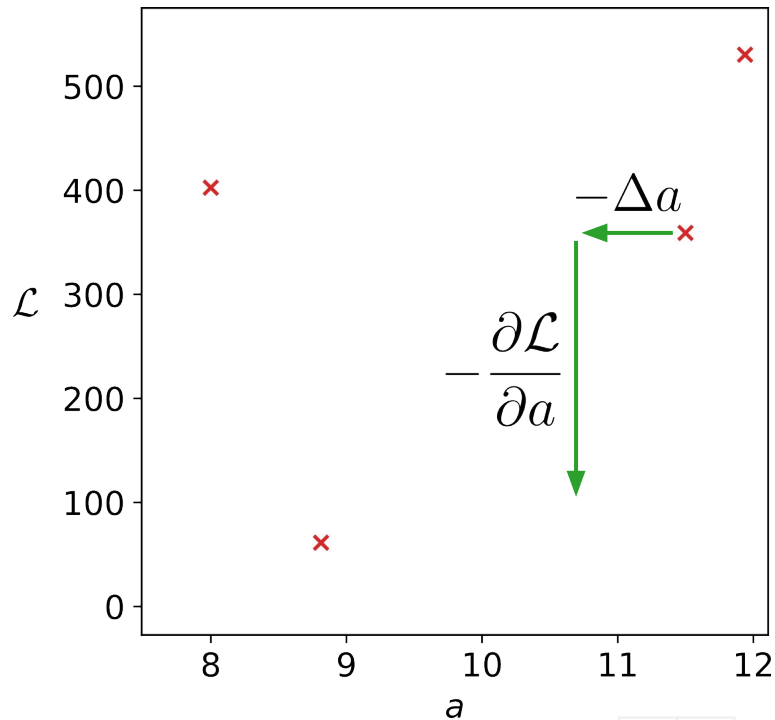
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

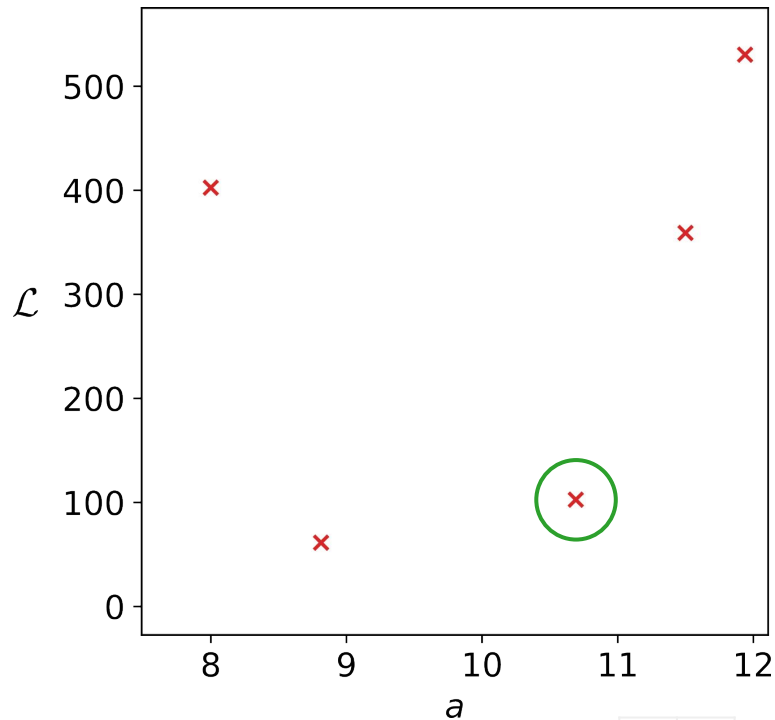
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

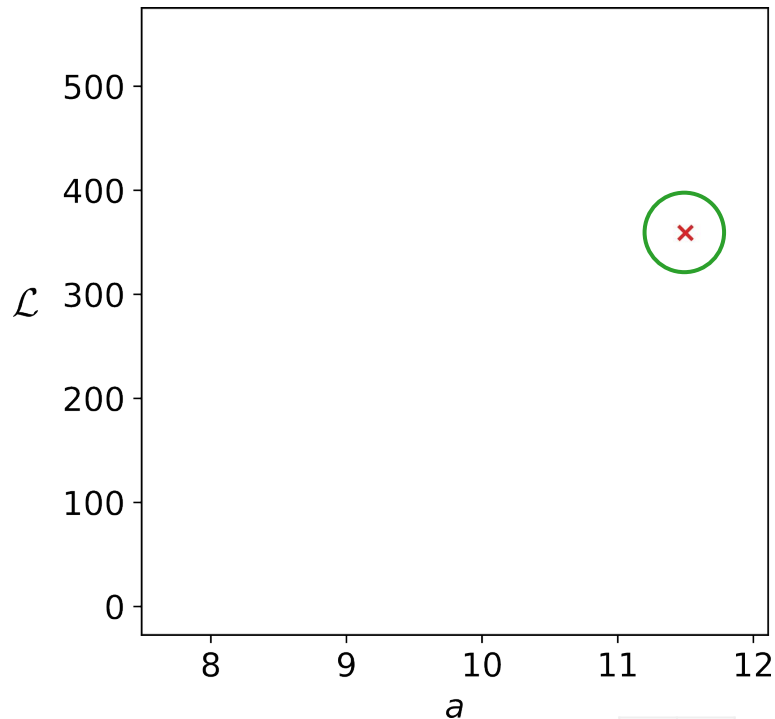
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

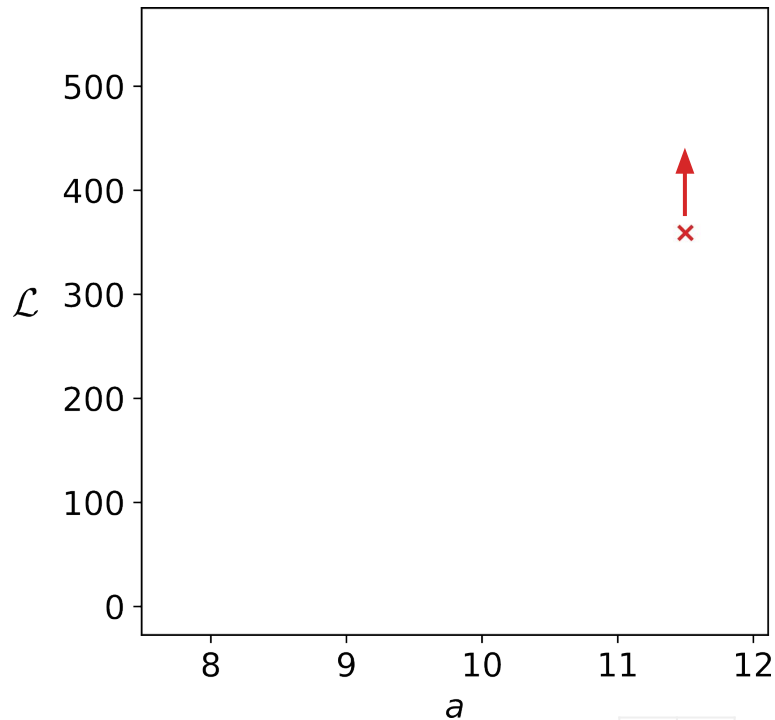
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

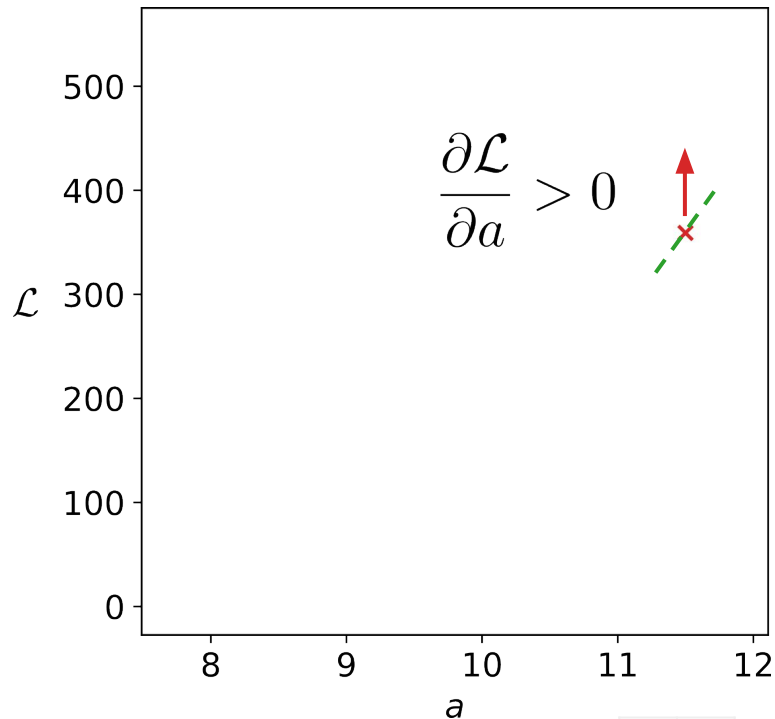
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

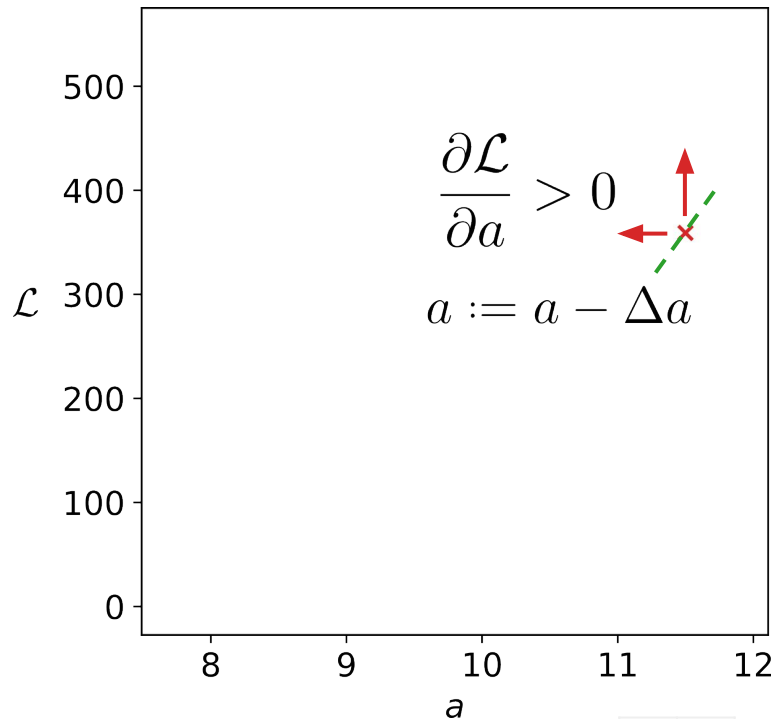
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“) \mathcal{L}

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

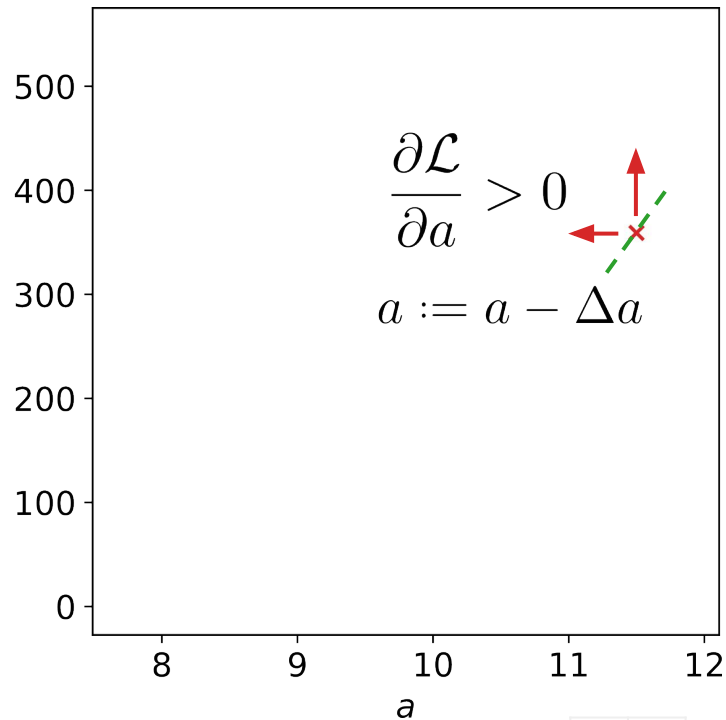
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$\frac{\partial \mathcal{L}}{\partial a} > 0$$

$$a := a - \Delta a$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“) \mathcal{L}

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

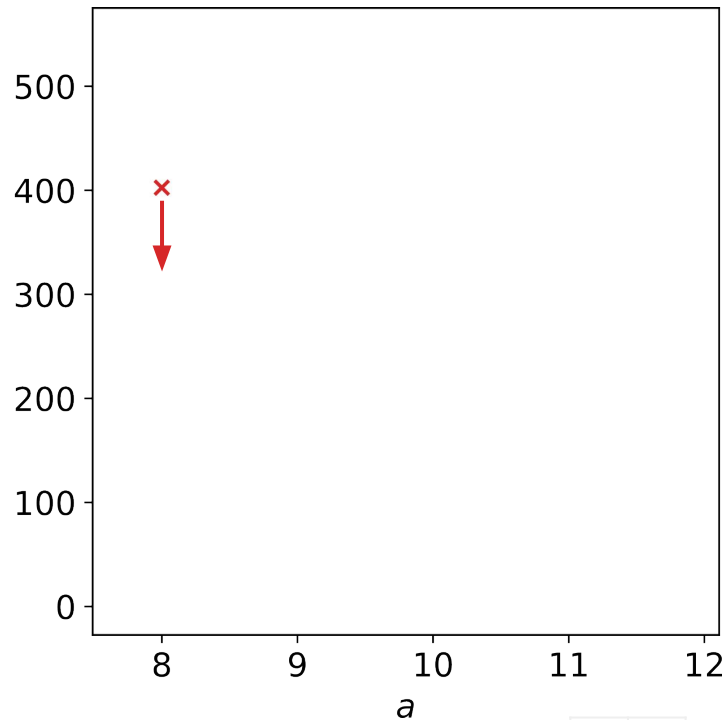
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$\frac{\partial \mathcal{L}}{\partial a} > 0$$

$$a := a - \Delta a$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

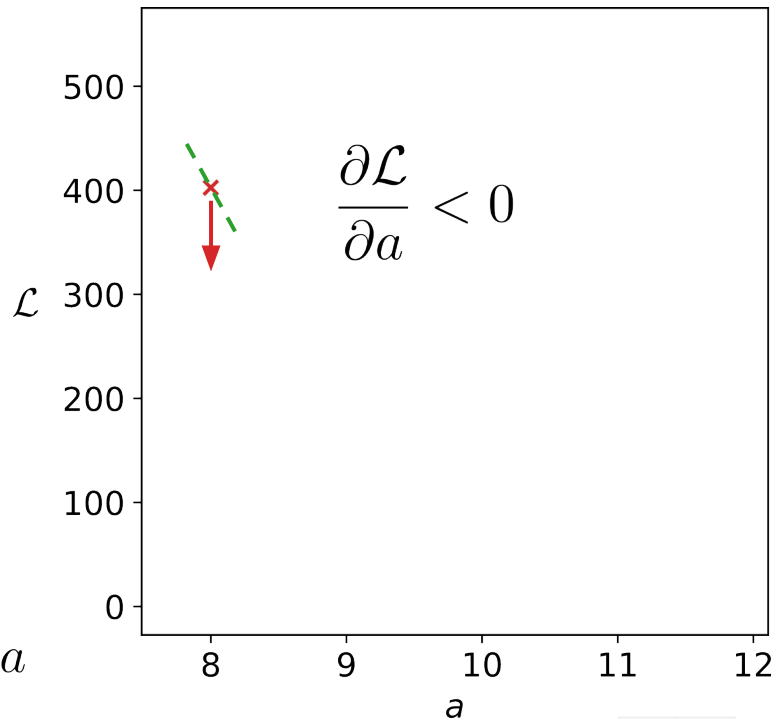
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$\frac{\partial \mathcal{L}}{\partial a} > 0$$

$$a := a - \Delta a$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a)$$

$$\Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

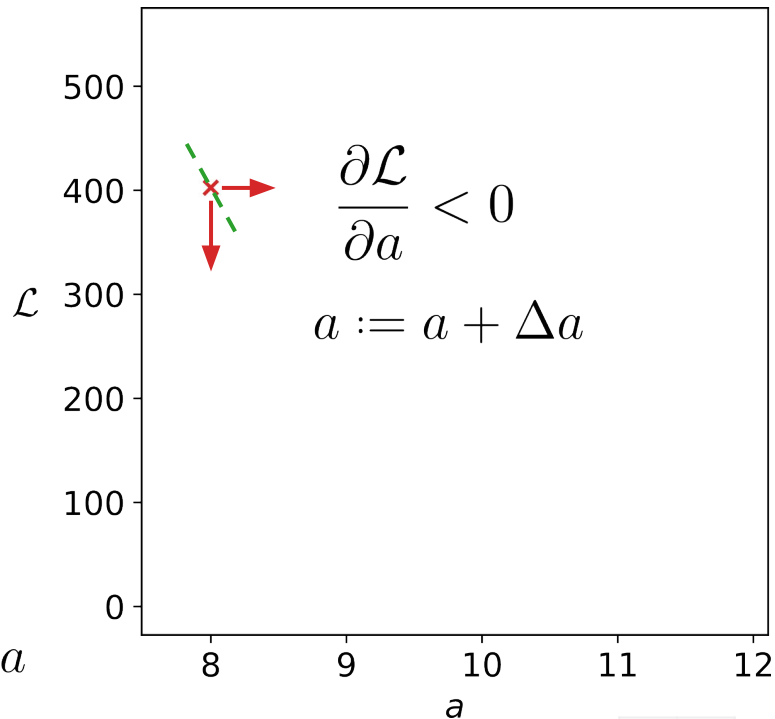
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$\frac{\partial \mathcal{L}}{\partial a} > 0$$

$$a := a - \Delta a$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

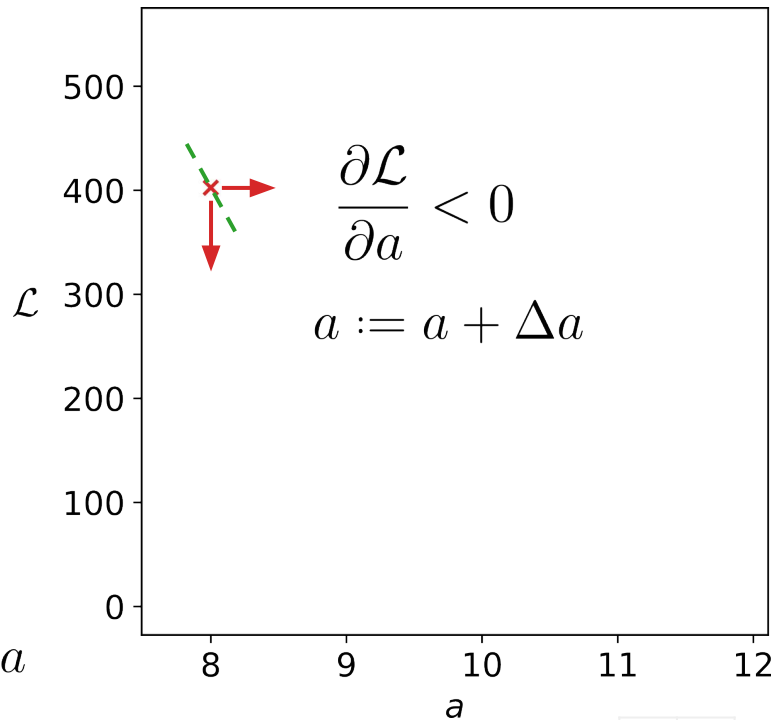
Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ a := a + \Delta a \end{array} \right| \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} > 0 \\ a := a - \Delta a \end{array}$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

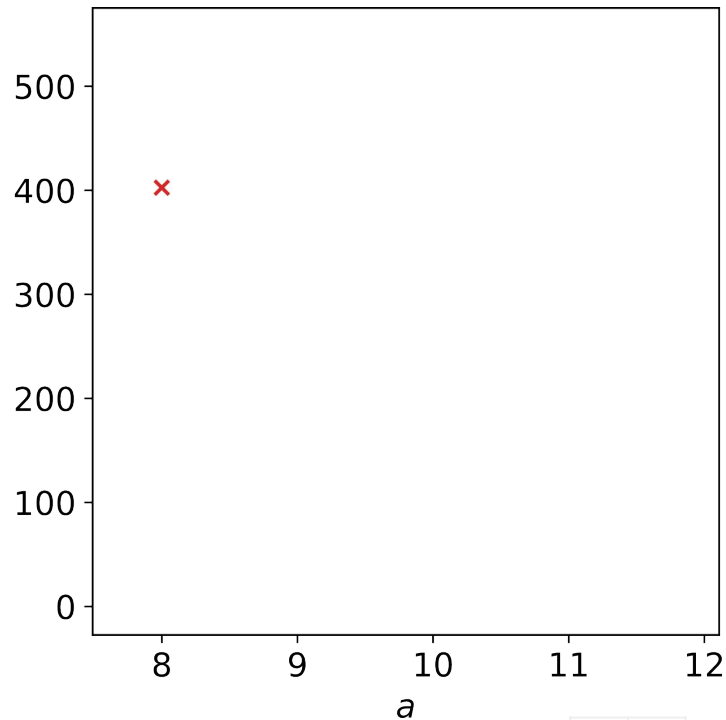
Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“) \mathcal{L}

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ a := a + \Delta a \end{array} \right| \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} > 0 \\ a := a - \Delta a \end{array}$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

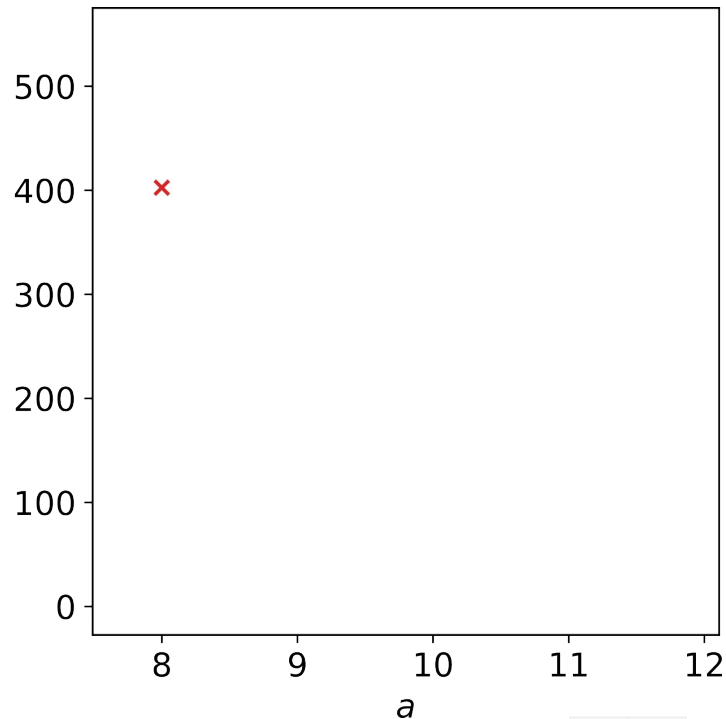
Išvestinė parodo funkcijos pokytį („įtaką“) \mathcal{L}

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ a := a + \Delta a \end{array} \right| \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} > 0 \\ a := a - \Delta a \end{array}$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \frac{\partial \mathcal{L}}{\partial a}$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

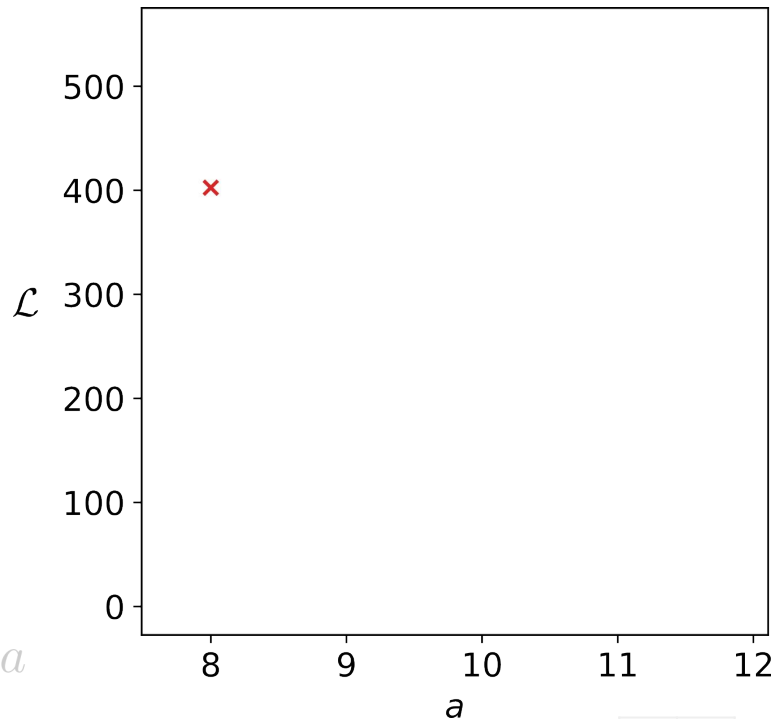
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right|$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad \left. \begin{array}{l} a := a + \Delta a \\ a := a - \Delta a \end{array} \right|$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

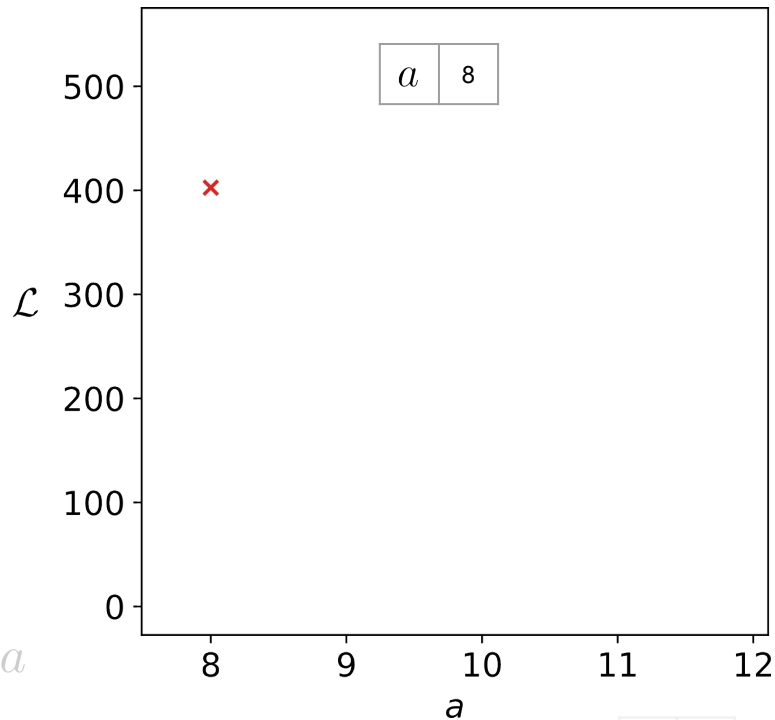
$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right|$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad a := a + \Delta a \quad \left| \quad a := a - \Delta a$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right| \begin{array}{l} a := a + \Delta a \\ a := a - \Delta a \end{array}$$
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

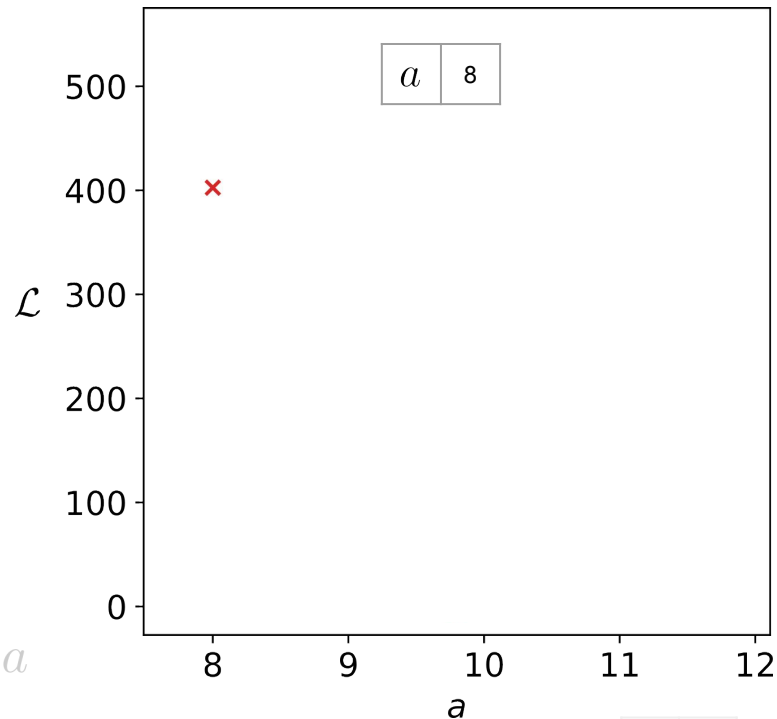
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right|$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad \left. \begin{array}{l} a := a + \Delta a \\ a := a - \Delta a \end{array} \right|$$

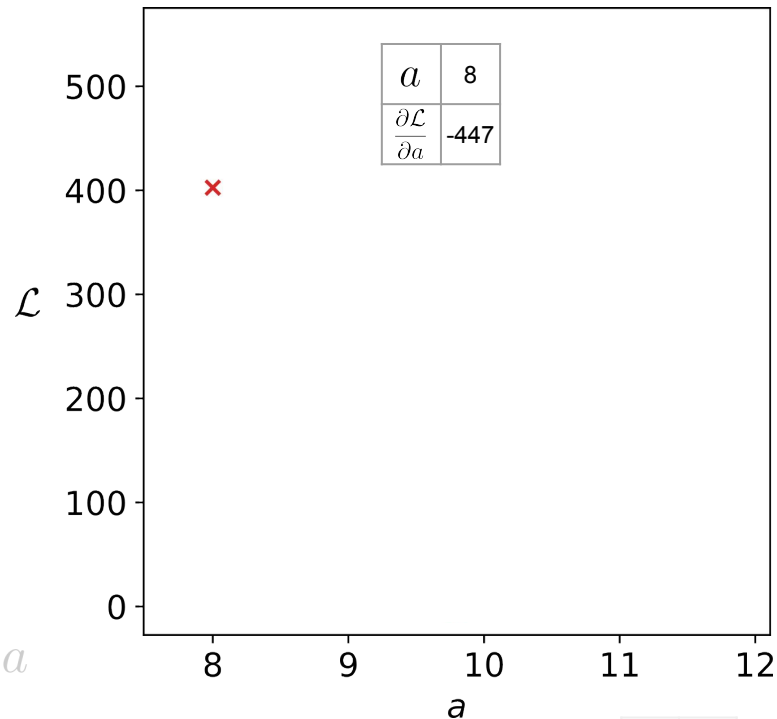
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} \approx -447 \end{array} \right| \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} > 0 \\ \frac{\partial \mathcal{L}}{\partial a} \approx 422 \end{array}$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad a := a + \Delta a \quad \left| \quad a := a - \Delta a$$

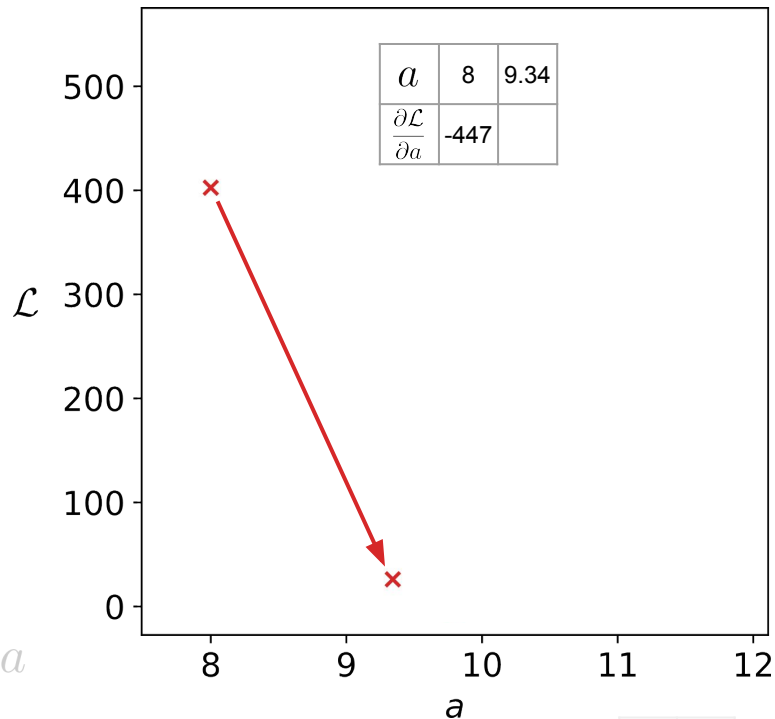
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} \approx -447 \end{array} \right| \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} > 0 \\ \frac{\partial \mathcal{L}}{\partial a} \approx 422 \end{array}$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad a := a + \Delta a \quad \left| \quad a := a - \Delta a$$

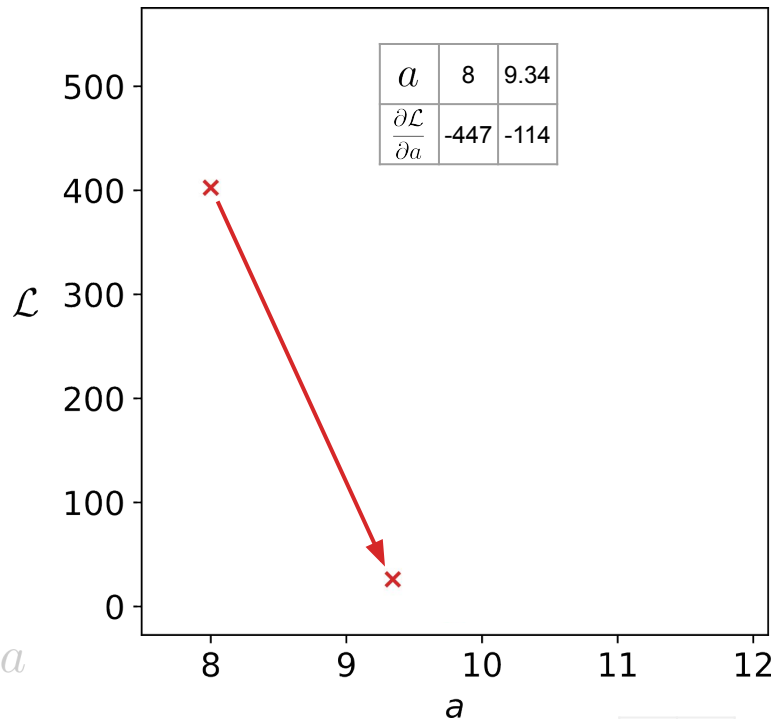
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right| \begin{array}{l} a := a + \Delta a \\ a := a - \Delta a \end{array}$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

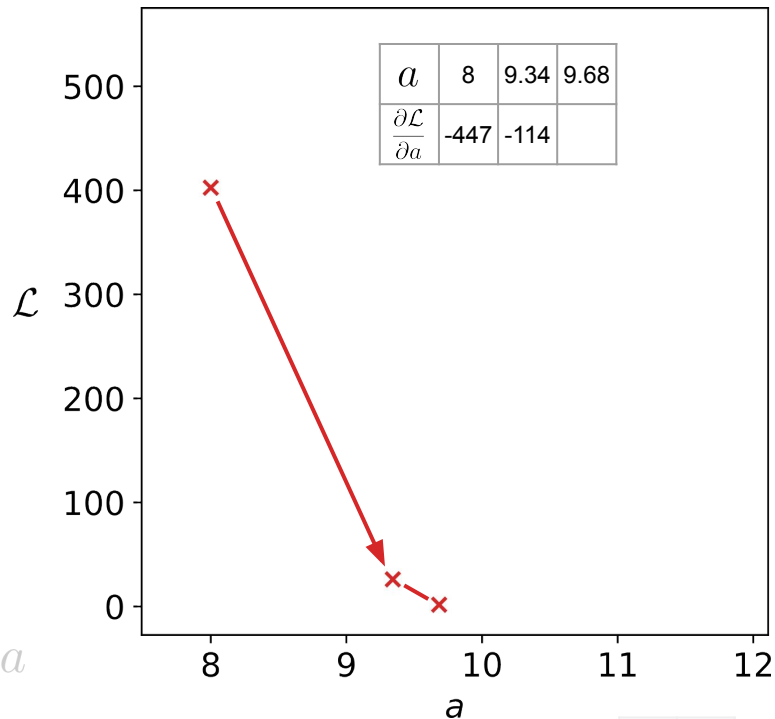
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right| \begin{array}{l} a := a + \Delta a \\ a := a - \Delta a \end{array}$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

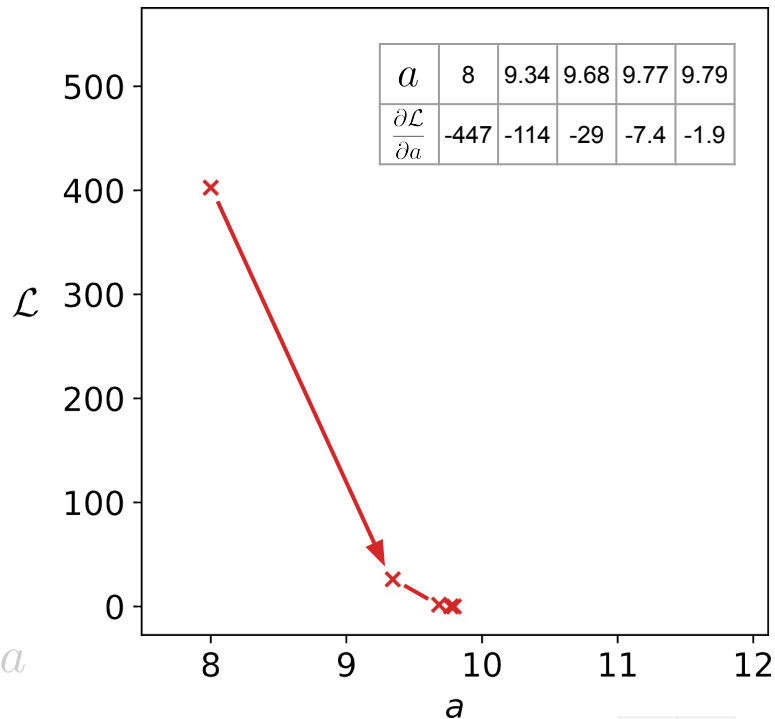
$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

γ parenkame patys

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$
$$\mathcal{L}(8) \approx 402 \quad \left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial a} < 0 \\ \frac{\partial \mathcal{L}}{\partial a} > 0 \end{array} \right|$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447 \quad a := a + \Delta a \quad \left| \quad a := a - \Delta a$$

$$\mathcal{L}(11.5) \approx 359$$

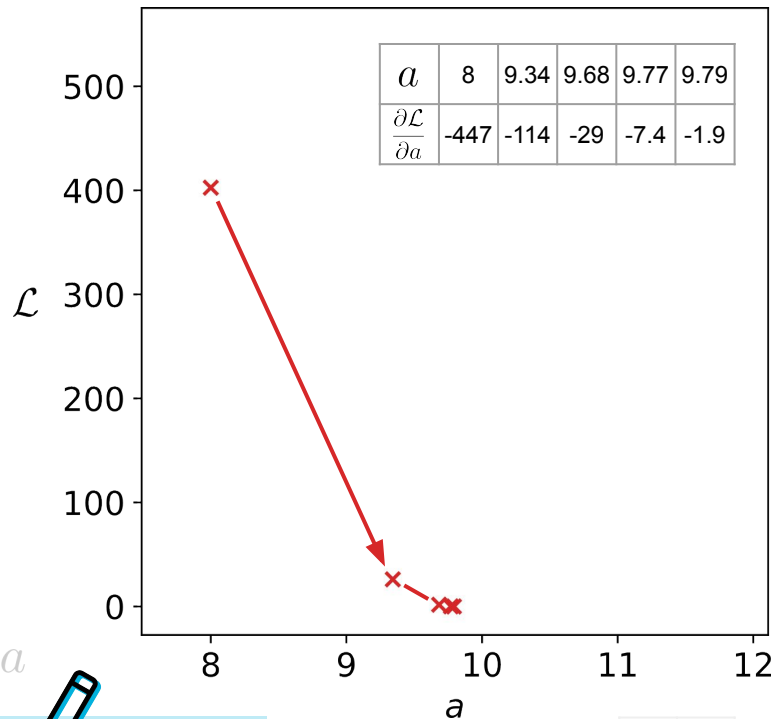
$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

kartoti:

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

kol pasieksime minimumą

$$\gamma = 0.003$$



$$\hat{y} = ax$$

$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

Kaip rasti minimumą? $\min_a \mathcal{L}$

Išvestinė parodo funkcijos pokytį („įtaką“)

$$\frac{\partial \mathcal{L}}{\partial a} \quad \mathcal{L} = \mathcal{L}(a) \quad \Delta a > 0$$

$$\mathcal{L}(8) \approx 402$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx -447$$

$$\mathcal{L}(11.5) \approx 359$$

$$\frac{\partial \mathcal{L}}{\partial a}(8) \approx 422$$

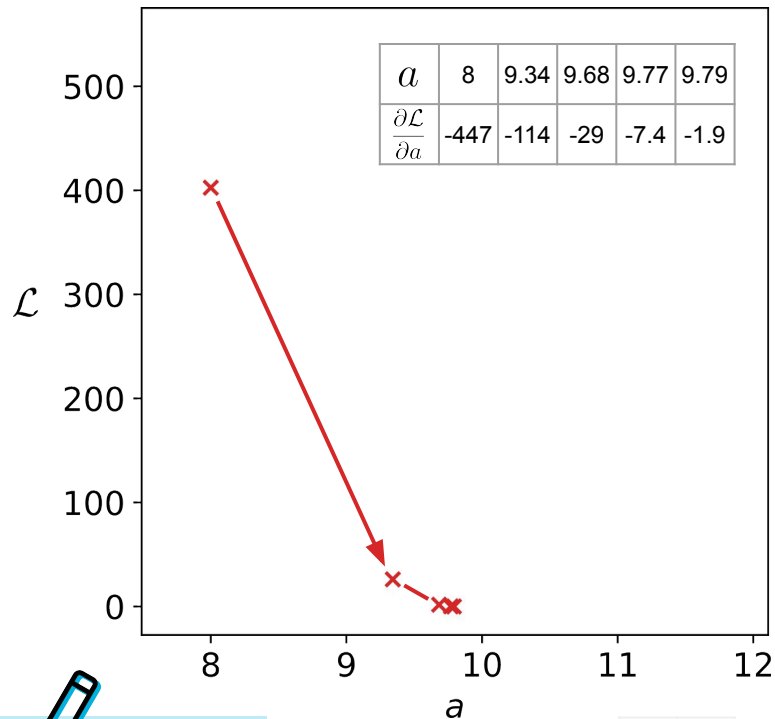
γ – learning rate

kartoti:

$$a := a - \gamma \frac{\partial \mathcal{L}}{\partial a}$$

kol pasieksime minimumą

$\gamma = 0.003$



$$\hat{y} = ax$$

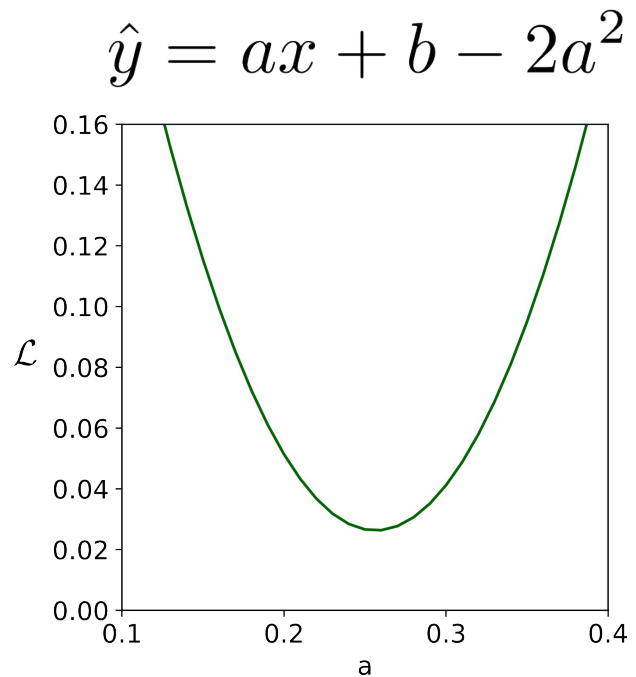
$$\mathcal{L} = \sum_i (\hat{y}_i - y_i)^2$$

x	y
2	19.6
4.5	44.1
10	98

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

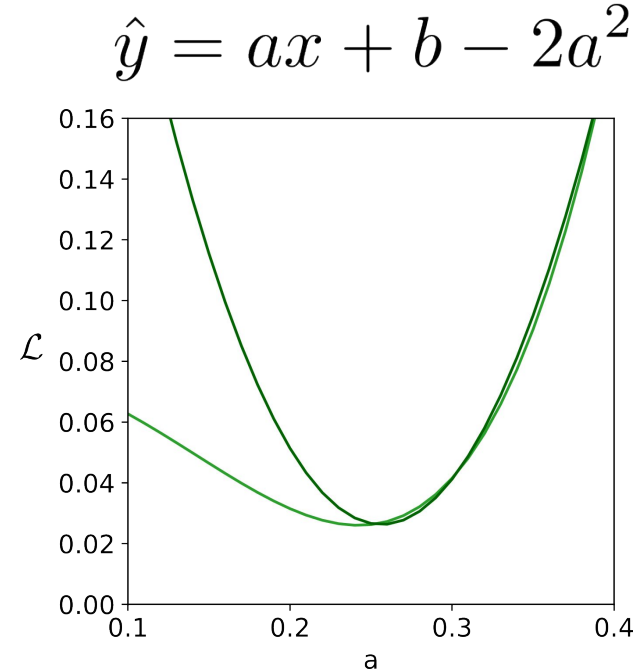
todo: gal kitai savaitei?



$b = 2$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

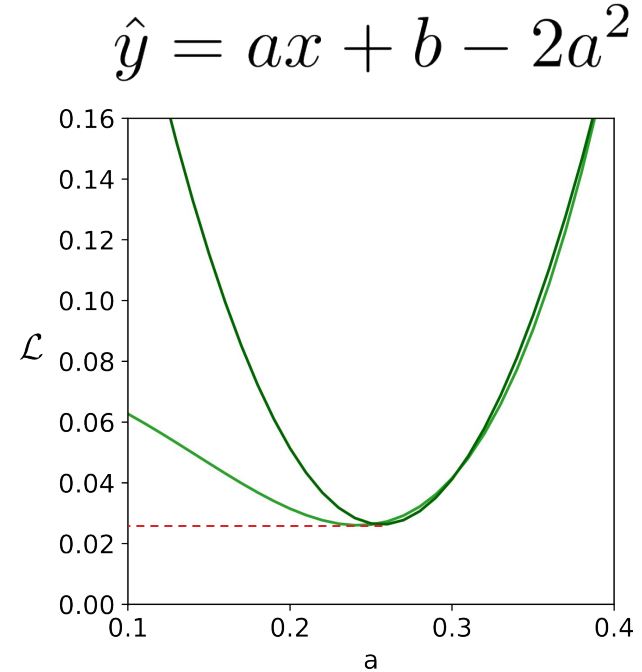


$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$$b = 2$$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85



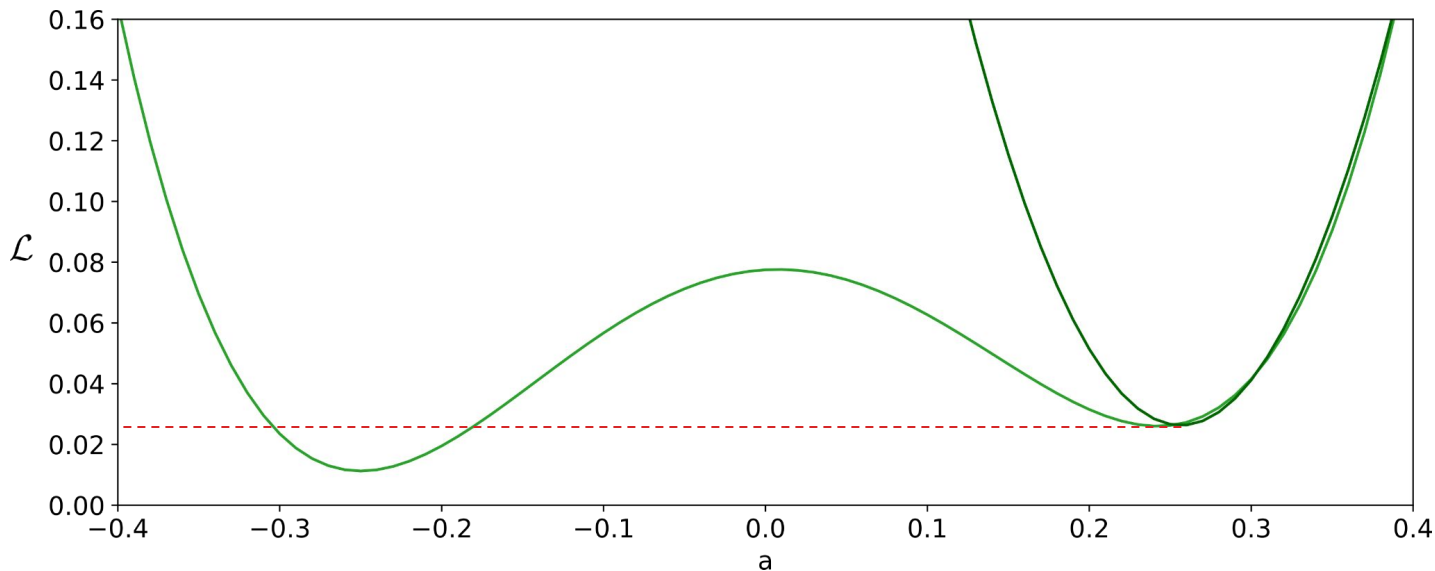
$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$$b = 2$$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

$$\hat{y} = ax + b - 2a^2$$



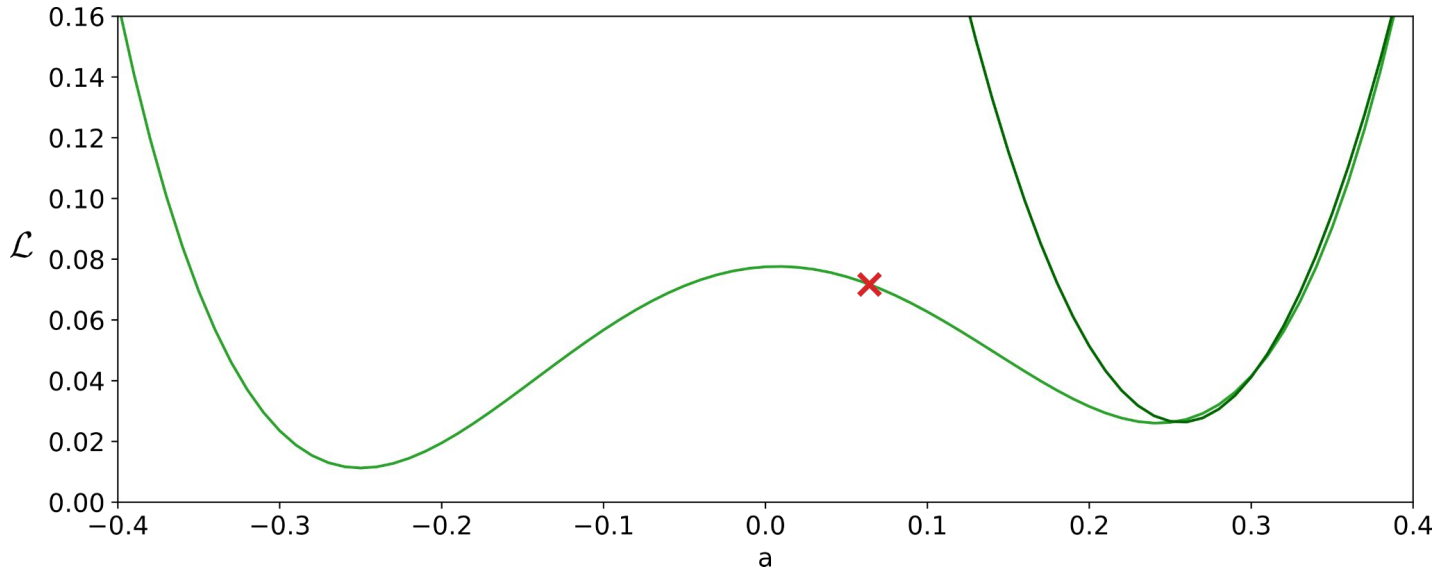
$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$$b = 2$$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

$$\hat{y} = ax + b - 2a^2$$



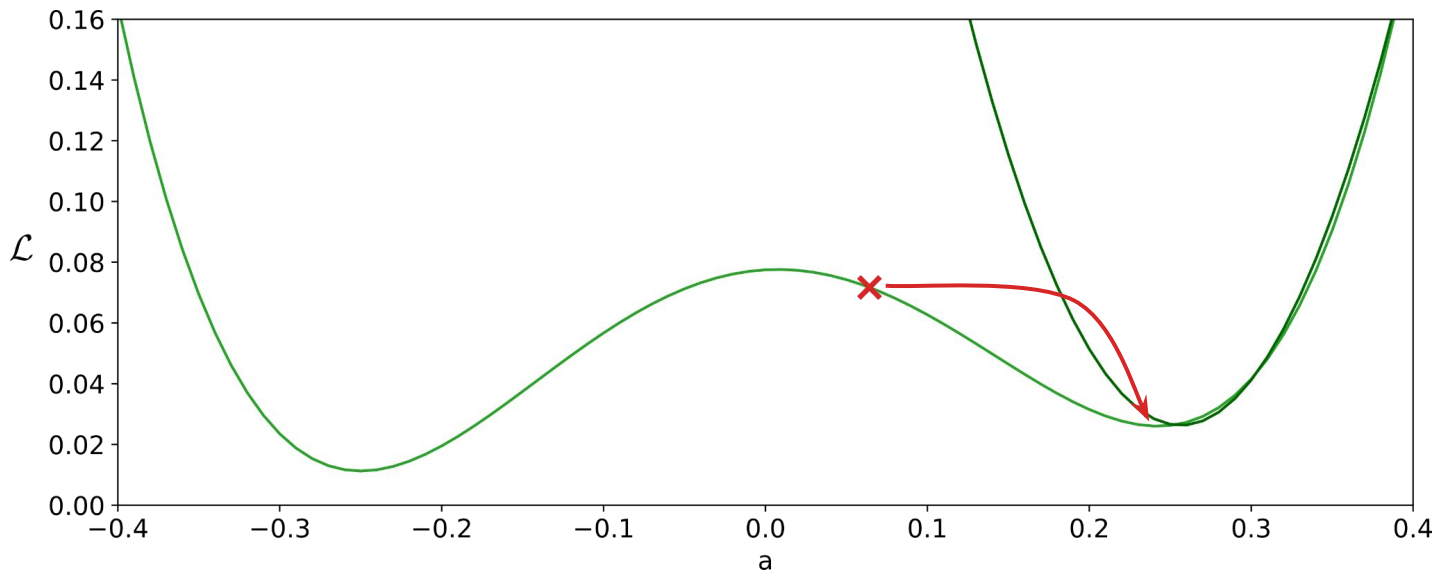
$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$$b = 2$$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

$$\hat{y} = ax + b - 2a^2$$



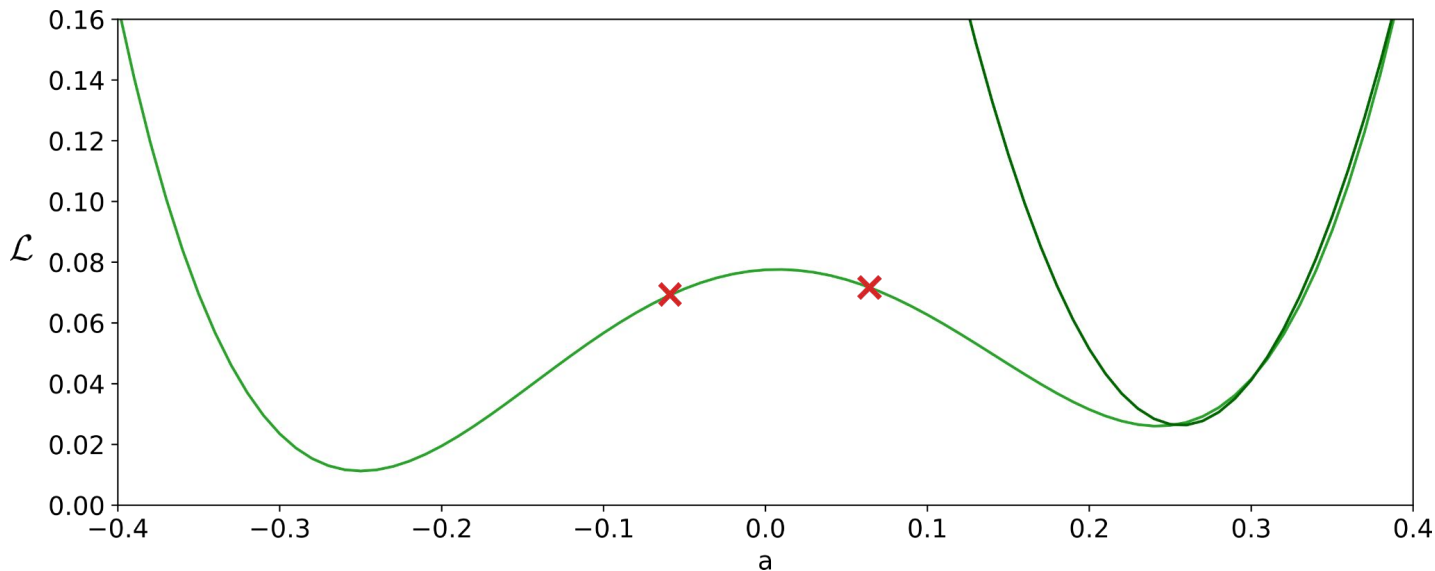
$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

b = 2

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

$$\hat{y} = ax + b - 2a^2$$



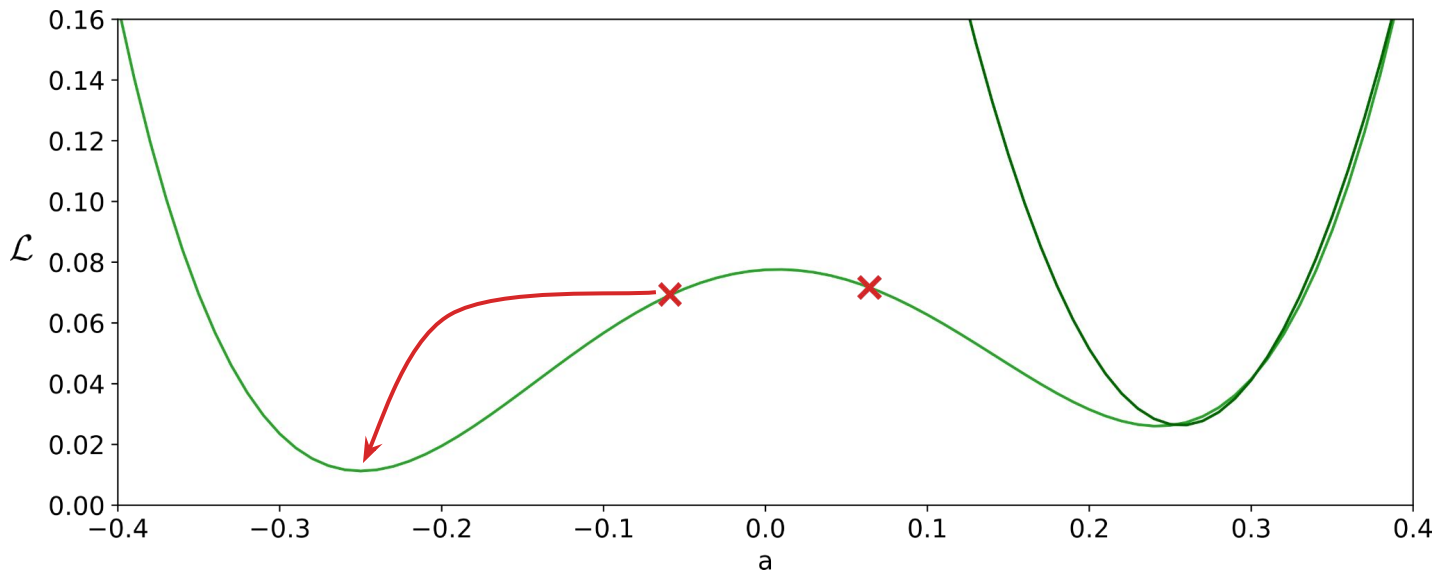
$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$$b = 2$$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

$$\hat{y} = ax + b - 2a^2$$

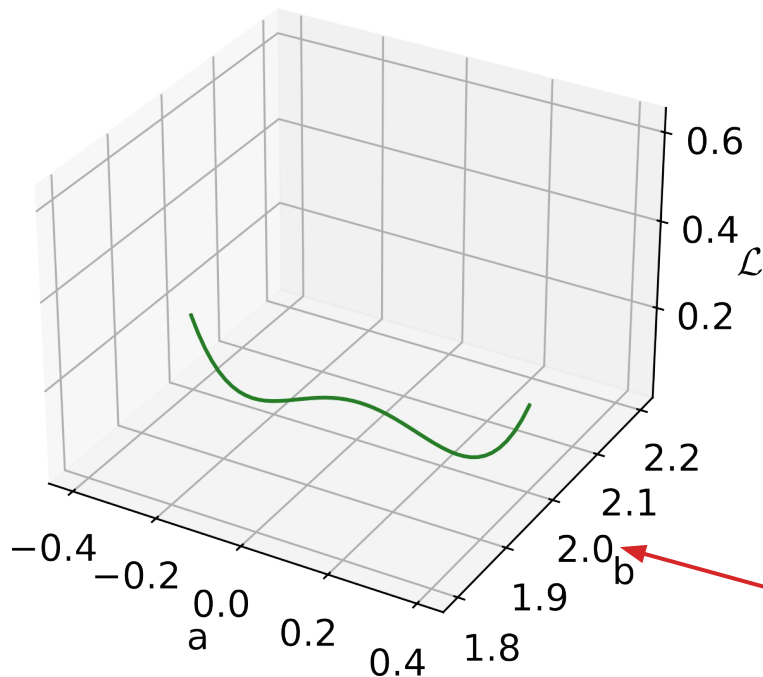


$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

b = 2

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85

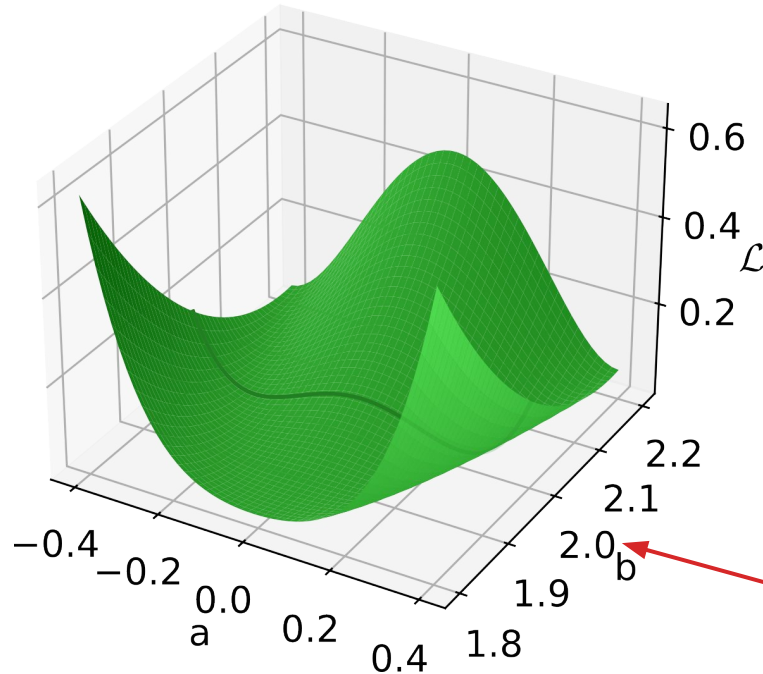


$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$b = 2$

Gradientinis nusileidimas

x	y
-0.3	1.9
-0.2	1.85
0.2	1.85
0.3	1.85



$$\hat{y} = -5.5(a - 0.246)x + b - 0.217$$

$b = 2$

Recap

 Duomenys vs. modelio spėjimai

 Kaip įvertinti modelio „gerumą“

 Kaip apmokyti modelį? (Kaip rasti tinkamą α ?)

Data vs. Predictions

L2 Loss Function

Gradient Descent



sasnauskas.net/gmm