Artificial Intelligence Foundation – JC3001

Lecture 39: Machine Learning – Regression II

Prof. Aladdin Ayesh (aladdin.ayesh@abdn.ac.uk)

Dr. Binod Bhattarai (binod.bhattarai@abdn.ac.uk)

Dr. Gideon Ogunniye, (g.ogunniye@abdn.ac.uk)

November 2025

Material adapted from:
Russell and Norvig (AIMA Book): Chapter 19 (19.4–19.6)
Sebastian Thrun (Stanford University / Udacity)
Andrew Ng (Stanford University / Coursera)

Course Progression

- Part 1: Introduction
  1. Introduction to AI ✓
  2. Agents ✓
- Part 2: Problem-solving
  1. Search 1: Uninformed Search ✓
  2. Search 2: Heuristic Search ✓
  3. Search 3: Local Search ✓
  4. Search 4: Adversarial Search ✓
- Part 3: Reasoning and Uncertainty
  1. Reasoning 1: Constraint Satisfaction ✓
  2. Reasoning 2: Logic and Inference ✓
  3. Probabilistic Reasoning 1: BNs ✓
  4. Probabilistic Reasoning 2: HMMs ✓
- Part 4: Planning
  1. Planning 1: Intro and Formalism ✓
  2. Planning 2: Algorithms & Heuristics ✓
  3. Planning 3: Hierarchical Planning ✓
  4. Planning 4: Stochastic Planning ✓
- Part 5: Learning
  1. Learning 1: Intro to ML ✓
  2. **Learning 2: Regression**
  3. Learning 3: Neural Networks
  4. Learning 4: Reinforcement Learning
- Part 6: Conclusion
  1. Ethical Issues in AI
  2. Conclusions and Discussion

Outline

- Cost Function
- Error Function

Cost Function—Intuition I

Hypothesis:

$h_{\boldsymbol{w}}(x^{(i)}) = \boldsymbol{w}_0 + \boldsymbol{w}_1 x^{(i)}$

Parameters:

$\boldsymbol{w}_0, \boldsymbol{w}_1$

Cost Function:

$$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)}\right)^2$$

Goal: $\min_{\boldsymbol{w}_0, \boldsymbol{w}_1} Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

Simplified

$h_{\boldsymbol{w}}(x^{(i)}) = \boldsymbol{w}_1 x^{(i)}$

$\boldsymbol{w}_1$

$$Loss(\boldsymbol{w}_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)}\right)^2$$

$\min_{\boldsymbol{w}_1} Loss(\boldsymbol{w}_1)$

Cost Function—Example I

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)

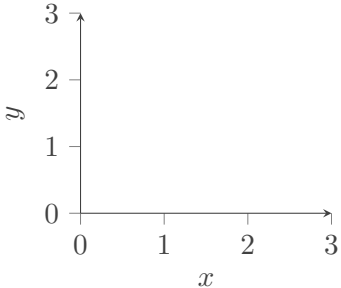
$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

Cost Function—Example I

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



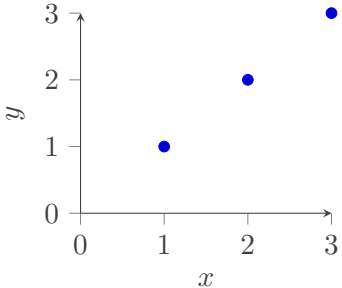
$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

Cost Function—Example I

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



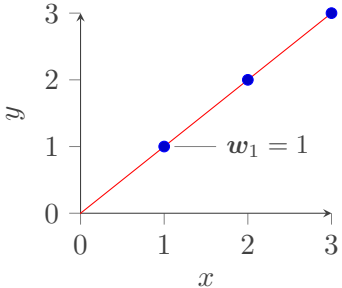
$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

Cost Function—Example I

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



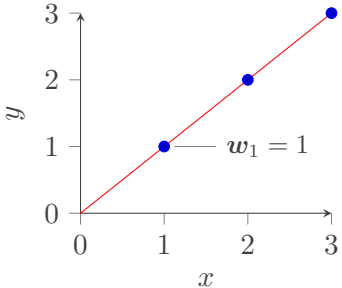
$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

Cost Function—Example I

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

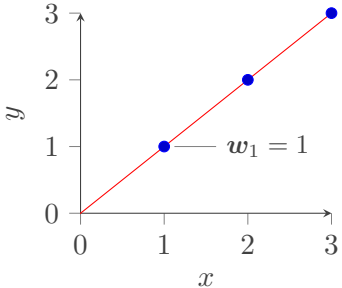
(function of the parameter $\boldsymbol{w}_1$)

$$
\begin{aligned}
Loss(\boldsymbol{w}_1) &= \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)} \right)^2 \\
&= \frac{1}{2m} \left(0^2 + 0^2 + 0^2 \right) \\
&= 0
\end{aligned}
$$

Cost Function—Example I

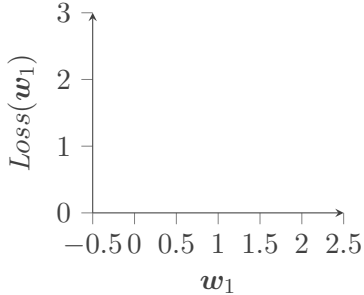
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

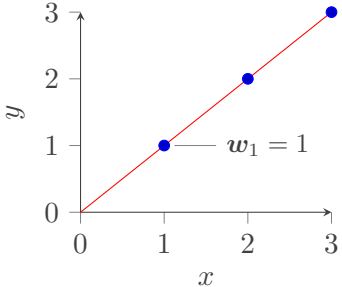


$$
\begin{aligned}
Loss(\boldsymbol{w}_1) &= \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)} \right)^2 \\
&= \frac{1}{2m} \left(0^2 + 0^2 + 0^2 \right) \\
&= 0
\end{aligned}
$$

Cost Function—Example I

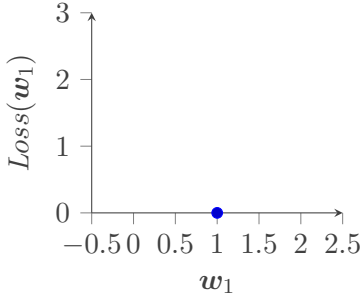
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

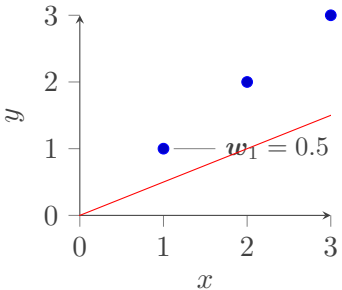


$$
\begin{aligned}
Loss(\boldsymbol{w}_1) &= \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)} \right)^2 \\
&= \frac{1}{2m} \left(0^2 + 0^2 + 0^2 \right) \\
&= 0
\end{aligned}
$$

Cost Function—Example II

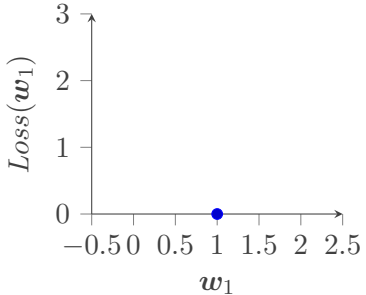
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

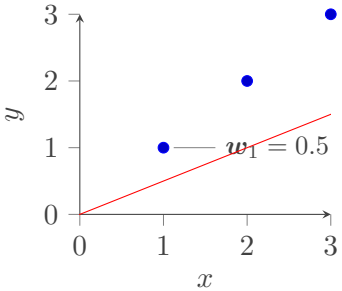
(function of the parameter $\boldsymbol{w}_1$)



Cost Function—Example II

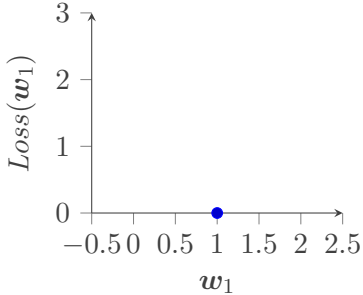
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

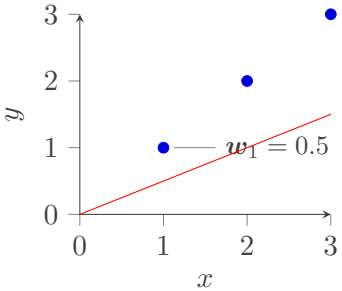


$$Loss(0.5) = \frac{1}{2m}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$
$$= \frac{1}{2*3}(3.5) = \frac{3.5}{6}$$
$$\approx 0.58$$

Cost Function—Example II

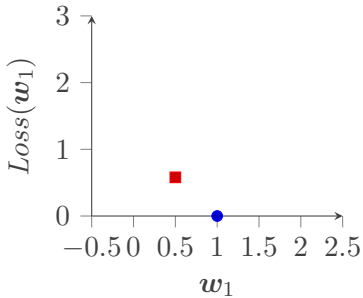
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)

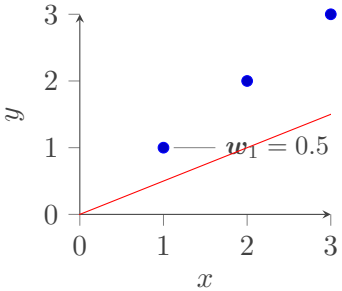


$$Loss(0.5) = \frac{1}{2m}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$
$$= \frac{1}{2*3}(3.5) = \frac{3.5}{6}$$
$$\approx 0.58$$

Cost Function—Example II

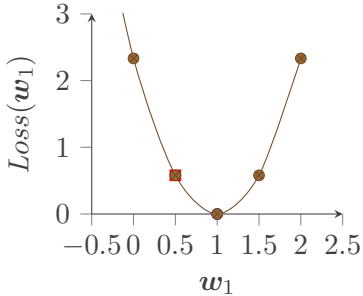
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_1)$

(function of the parameter $\boldsymbol{w}_1$)



$$Loss(0.5) = \frac{1}{2m}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$
$$= \frac{1}{2*3}(3.5) = \frac{3.5}{6}$$
$$\approx 0.58$$

Hypothesis: $h_{\boldsymbol{w}}(x^{(i)}) = \boldsymbol{w}_0 + \boldsymbol{w}_1 x^{(i)}$

Parameters: $\boldsymbol{w}_0, \boldsymbol{w}_1$

Cost Function: $Loss(\boldsymbol{w}_0, \boldsymbol{w}_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_{\boldsymbol{w}}(x^{(i)}) - y^{(i)} \right)^2$

Goal: $\min_{\boldsymbol{w}_0, \boldsymbol{w}_1} Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

Cost Function—Example III

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)

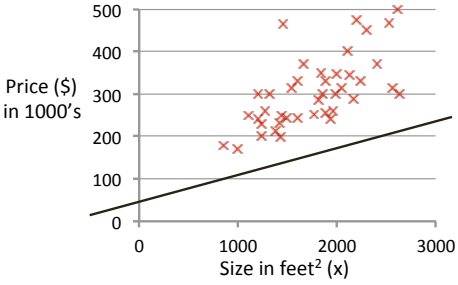
$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)

Cost Function—Example III

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



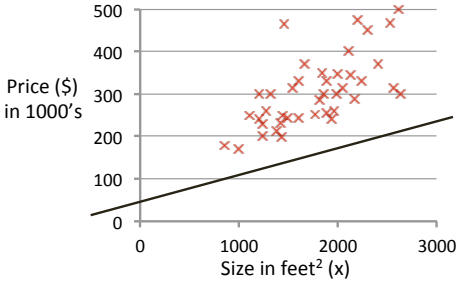
$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)

Cost Function—Example III

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



$$h_{\boldsymbol{w}}(x) = 50 + 0.06x$$

$$\boldsymbol{w}_0 = 50$$
$$\boldsymbol{w}_1 = 0.06$$

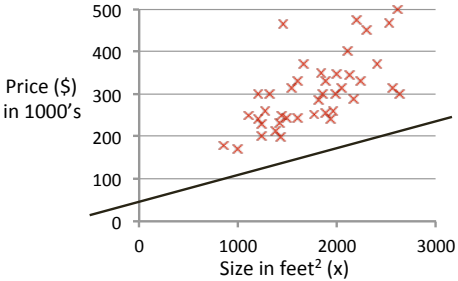
$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)

Cost Function—Example III

$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



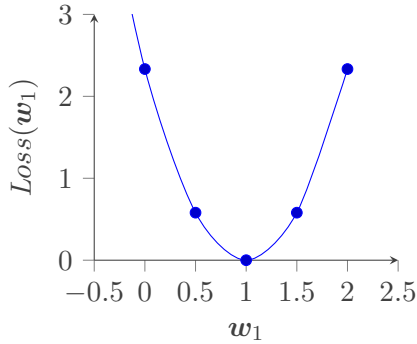
$h_{\boldsymbol{w}}(x) = 50 + 0.06x$

$\boldsymbol{w}_0 = 50$

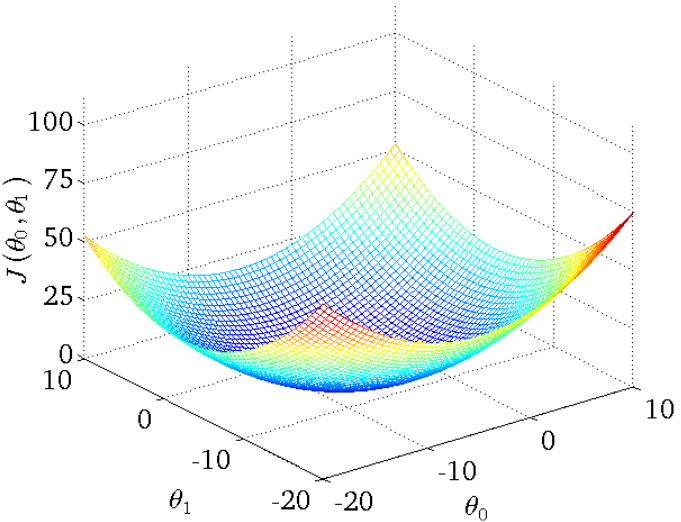
$\boldsymbol{w}_1 = 0.06$

$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)



This is how the cost function looked when we had only one $\boldsymbol{w}$.


100
75
50
25
0
$J(\theta_0, \theta_1)$
10
0
-10
-20
θ_1
-20
-10
0
10
θ_0

Outline

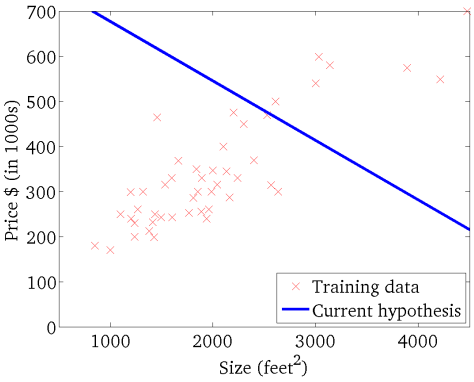
- Cost Function
- Error Function

Error Function I

2 Error Function

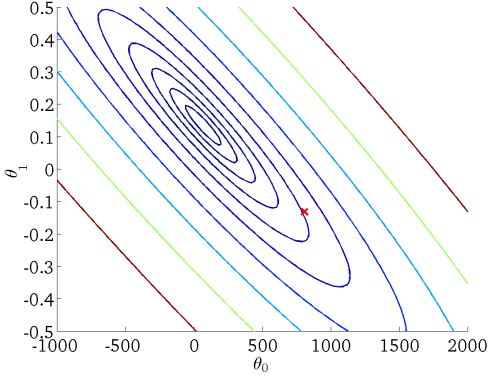
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)

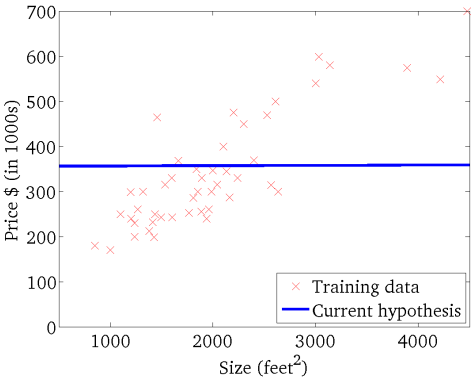


Contour plot

Error Function II

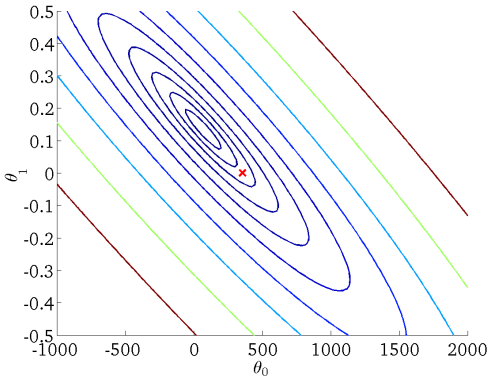
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

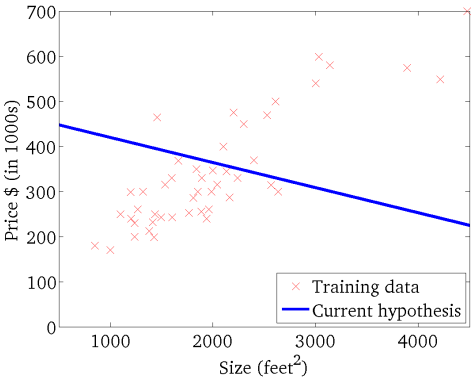
(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)



Error Function III

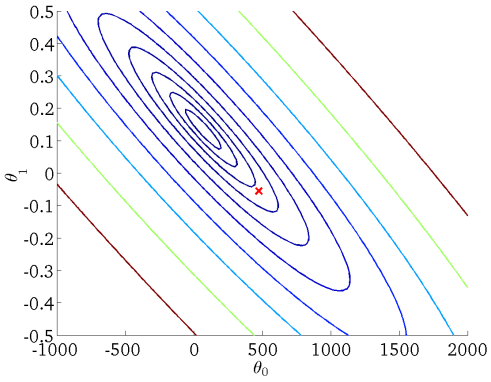
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

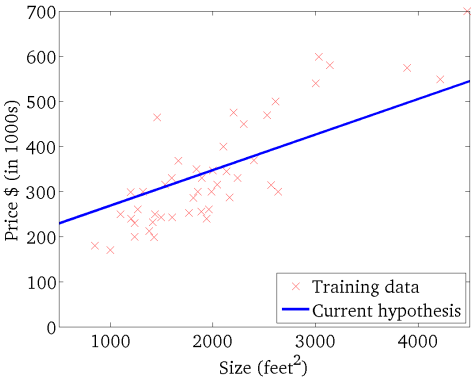
(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)



Error Function IV

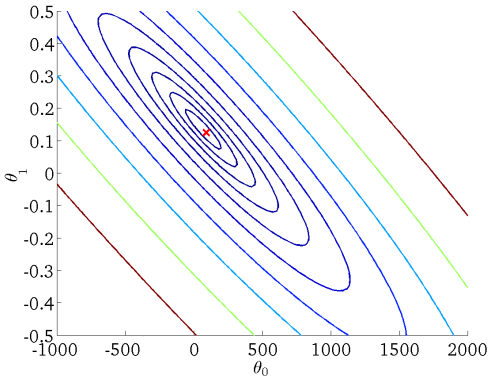
$h_{\boldsymbol{w}}(x)$

(for fixed $\boldsymbol{w}_0, \boldsymbol{w}_1$, this is a function of x)



$Loss(\boldsymbol{w}_0, \boldsymbol{w}_1)$

(function of the parameters $\boldsymbol{w}_0, \boldsymbol{w}_1$)



To continue in the next session.