Weekly Problem Set: Deep and Reinforcement Learning

Module 1, Week 1: The Learning Problem and Feedforward Networks

Srikanth Pai, Asst. Professor, MSE

Instructions

Answers should be concise and mathematically precise. State any theorem or result you use before applying it. This set covers both lectures of the week: Part A (Lecture 1.1, The Learning Problem) and Part B (Lecture 1.2, Feedforward Networks). DS students: Problems A4 and B4 require coded implementations in Python (NumPy only, no scikit-learn, no PyTorch, no TensorFlow); submit your source code together with a short write-up (half a page each) describing your observations. MBA students: Part C is a case-study writing exercise; submit a written report of two to three pages.

Part A: The Learning Problem

A1. Prove or disprove: if a learning algorithm achieves zero empirical risk, i.e. $\widehat{R}(\hat{h}) = 0$, then the true risk $R(\hat{h})$ is also zero.

Give a concrete counterexample if the statement is false: specify an explicit distribution P on $\mathcal{X} \times \mathcal{Y}$ (it can be supported on just two or three points), an explicit hypothesis class $\mathcal{H}$ containing at least two hypotheses, and a loss function L, such that the ERM solution satisfies $\widehat{R}(\hat{h}) = 0$ but $R(\hat{h}) > 0$.

A2. Prove the bias-variance decomposition stated in Lecture 1 (Theorem 3.1 of the notes). Specifically, suppose $y = f(x) + \varepsilon$ with $\mathbb{E}[\varepsilon] = 0$, $\mathrm{Var}(\varepsilon) = \sigma^2$, and ε independent of $\hat{h}(x)$. Write $\bar{h}(x) = \mathbb{E}[\hat{h}(x)]$. Show that

$$\mathbb{E}\big[(\hat{h}(x) - y)^2\big] \;=\; (\bar{h}(x) - f(x))^2 \;+\; \mathrm{Var}(\hat{h}(x)) \;+\; \sigma^2.$$

Your proof should identify exactly where each of the three terms comes from and why the cross terms vanish.

A3. Consider the hypothesis class $\mathcal{H}_d$ of univariate polynomials of degree at most d: $\mathcal{H}_d = \{h(x) = a_0 + a_1 x + \cdots + a_d x^d : a_i \in \mathbb{R}\}$. Suppose the true regression function is $f(x) = \sin(\pi x)$ on $[0, 1]$, with noise $\varepsilon \sim N(0, \sigma^2)$ for some small $\sigma > 0$.

(a) For $d = 1$, argue qualitatively (without computing) whether the bias or the variance is the dominant source of error at $x = 0.5$.

(b) What happens to the bias and variance as d grows from 1 to $n - 1$, where n is the number of training examples? Give a qualitative description for each regime: $d \ll n$, $d \approx n/2$, and $d = n - 1$.

(c) Suppose we add a second copy of every training example (so the dataset has $2n$ points, each repeated twice). Does this change the ERM solution $\hat{h}$ in $\mathcal{H}_d$ for any fixed d? Prove your answer.

A4. (DS students.) Implement polynomial regression from scratch using only NumPy. Apply it to the following synthetic dataset: generate $n = 30$ training points and $n_{\text{test}} = 200$ test points as

$$x_i \sim \text{Uniform}[0, 1], \qquad y_i = \sin(\pi x_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, 0.05^2).$$

For each polynomial degree $d \in \{1, 2, 3, 5, 8, 12, 20\}$:

(a) Fit the polynomial to the training data using the normal equations (compute the Vandermonde matrix $\Phi \in \mathbb{R}^{n\times(d+1)}$ and solve $\hat{a} = (\Phi^\top \Phi)^{-1}\Phi^\top y$).
(b) Compute and record the training MSE and the test MSE.

Plot both curves on the same axes (training MSE vs. d and test MSE vs. d).

Report: (i) the degree that minimises test MSE; (ii) the degree at which the training MSE first reaches below 0.001; (iii) a one-paragraph explanation of what you observe, using the vocabulary of bias, variance, underfitting, and overfitting.

[Hint: for numerical stability, use `numpy.linalg.lstsq` instead of inverting $\Phi^\top \Phi$ directly when d is large.]

Part B: Feedforward Networks

B1. Let $\mathcal{X} = \{[0, 0]^\top, [0, 1]^\top, [1, 0]^\top, [1, 1]^\top\}$ with XOR labels $\{0, 1, 1, 0\}$. Let $f(\mathbf{x}) = w_1 x_1 + w_2 x_2 + b$.

(a) Compute $J(w_1, w_2, b) = \frac{1}{4}\sum_{\mathbf{x}\in\mathcal{X}}(f^*(\mathbf{x}) - f(\mathbf{x}))^2$ explicitly as a polynomial in w_1, w_2, b.
(b) Set $\partial J/\partial w_1 = \partial J/\partial w_2 = \partial J/\partial b = 0$ and solve the resulting system.
(c) Verify that the minimiser gives $f(\mathbf{x}) = \frac{1}{2}$ for all $\mathbf{x} \in \mathcal{X}$.

B2. For $\ell = 1, \ldots, k$, let $f^{(\ell)}(\mathbf{x}) = A_\ell \mathbf{x} + \mathbf{b}_\ell$ where A_ℓ is a matrix and $\mathbf{b}_\ell$ a vector of appropriate dimensions.

(a) Prove by induction on k that the composition $F_k = f^{(k)} \circ \cdots \circ f^{(1)}$ is an affine function of the input; that is, $F_k(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ for some matrix A and vector $\mathbf{b}$.
(b) Conclude that no affine network of any depth can represent XOR exactly. Which theorem from the notes justifies this conclusion?

B3. (Preparation for Lecture 1.3.) Consider a single-hidden-unit network with one input $x \in \mathbb{R}$:

$$h = \max(0,\, w_1 x + c), \qquad \hat{y} = w_2 h + b, \qquad L = \tfrac{1}{2}(y - \hat{y})^2.$$

(a) For the case $w_1 x + c > 0$, compute $\partial L/\partial w_1$ and $\partial L/\partial c$ by applying the chain rule step by step. Write out every intermediate derivative explicitly.
(b) For the case $w_1 x + c < 0$, what are $\partial L/\partial w_1$ and $\partial L/\partial c$?
(c) Explain in one sentence why the non-differentiability of ReLU at $z = 0$ causes no practical difficulty during training.

B4. (DS students.) Implement the XOR network from the notes in Python using only NumPy. Use the parameters:

$$W = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad b = 0.$$

(a) For each of the four inputs $\mathbf{x} \in \mathcal{X}$, compute the hidden representation $\mathbf{h} = \max(0,\, W^\top \mathbf{x} + \mathbf{c})$ and the output y. Print a table matching the verification in the notes.

(b) Scatter-plot the four inputs in the (h_1, h_2) plane, using different markers or colours for XOR label 0 and label 1. Note: two inputs map to the same point in h-space. Mark this point with the label of both (they agree).

(c) On the same plot, draw the line $h_1 - 2h_2 = \frac{1}{2}$ (the decision boundary corresponding to $\mathbf{w}^\top \mathbf{h} = \frac{1}{2}$). Verify visually that it separates the two classes.

(d) Now change c_2 from -1 to 0 (so $\mathbf{c} = [0, 0]^\top$). Recompute the outputs for all four inputs. Which entries are now wrong? Explain in one sentence why this change breaks the network.

Part C: Case Study (MBA students)

C1. Zillow Offers and the price of a model that only looked accurate. In 2021 the US real-estate firm Zillow shut down its iBuying arm, Zillow Offers, took a write-down of roughly $500 million, and cut about a quarter of its workforce. The business bought homes automatically at prices set by a machine learning valuation model (the "Zestimate") and resold them at a small margin. The model performed well on historical data, yet the unit lost money at scale and was closed within months.

Write a case-study report of two to three pages addressing the following.

(a) **Background.** Summarise what Zillow Offers did, how the valuation model was used to make purchase decisions, and the timeline of the shutdown. Cite your sources.

(b) **Empirical risk versus true risk.** Zillow's model reportedly had a low median error on past sales. Using the vocabulary of this week's lectures, explain the distinction between the error measured on historical data ($\widehat{R}$) and the error the firm actually incurred when trading on the model's predictions (R). Why can the first be small while the second is large?

(c) **Distribution shift.** A central technical claim is that the housing market in 2021 moved away from the conditions the model was trained on. Explain what "distribution shift" means here and why a model with low in-sample error can fail badly when the deployment distribution P_{deploy} differs from the training distribution P_{train}. Relate this to Problem A1.

(d) **The feedback loop.** Unlike a passive forecast, Zillow's predictions caused purchases, which in turn affected local prices and the data the model later saw. Argue why acting on a prediction changes the very distribution the prediction assumed. What does this imply for evaluating a model by backtesting alone?

(e) **Recommendation.** As an analyst advising the firm before launch, propose two concrete safeguards (technical, organisational, or financial) that would have reduced the downside. For each, state the assumption it protects against and its cost.

Grading rewards correct use of the week's concepts (empirical vs. true risk, generalisation, overfitting, distribution shift) over narrative length.