

Submission format

Submit a PDF-version report for all problems, including source codes and print screen results. You can use any text/code editor. Google Colab is recommended.

Review lecture slide 4 and 5 and finish the following problems.

Problem 1 (Back-propagation)

Use back-propagation to calculate the gradients of

$$f(\mathbf{W}, \mathbf{x}) = \|\sigma(\mathbf{W}\mathbf{x})\|^2$$

with respect to \mathbf{x} and \mathbf{W} . Here, $\|\cdot\|^2$ is the calculation of L2 loss, \mathbf{W} is a 3×3 matrix, and \mathbf{x} is a 3×1 vector, and $\sigma(\cdot)$ is the ReLU function that performs **element-wise** operation.

Problem 2 (Linear Classifier using GD)

In this problem, you need to use Gradient Descent (GD) to train the linear classifier in the HW1, i.e., find the parameters \mathbf{W} , and then use it to recognize handwritten digits. Adopt still "Cross Entropy" as the loss function.

Requirements: 1) manually derive the gradients of linear classifier when using cross-entropy as the loss function, and write codes to implement it in recognizing handwritten digits; 2) the test accuracy should be at least 85%.