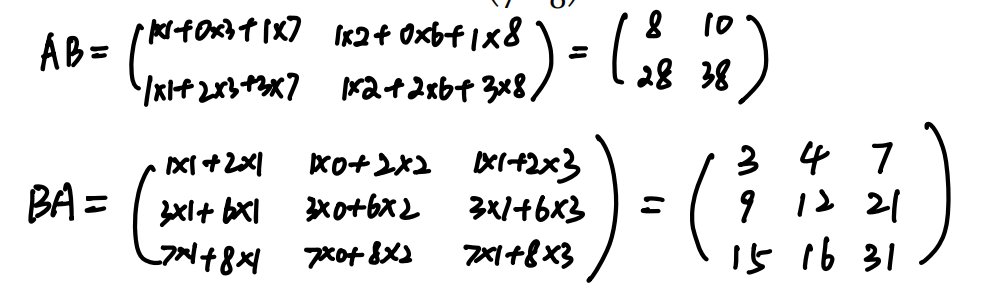
# 11040A Neural Networks

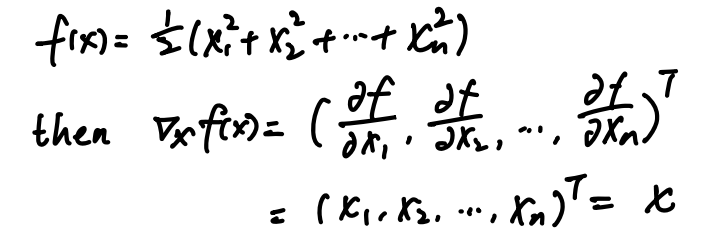
**Homework 1**

**Due Date: May 1, 2025 24:00**

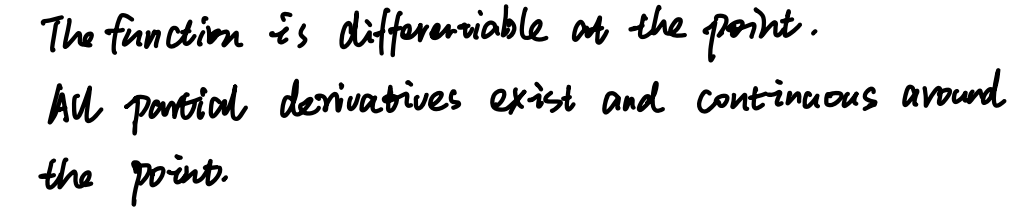
1. Let and . Compute and .



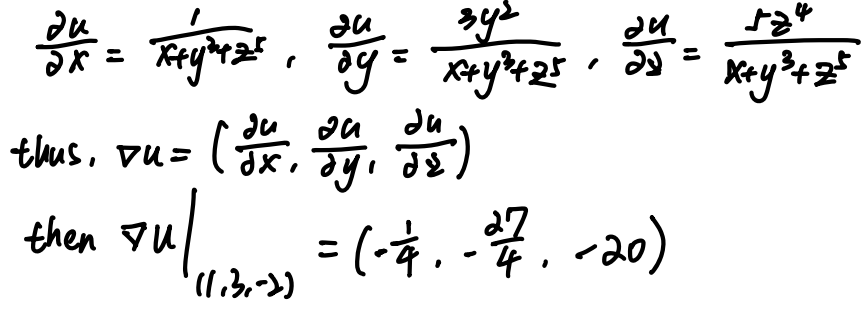
1. Let be a column vector, . Compute .



1. What are the conditions for a function to have a gradient vector at a point?

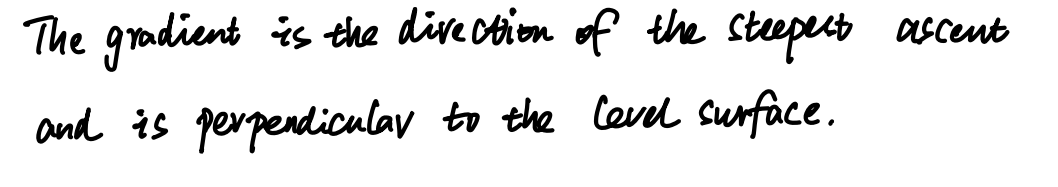


1. Let . Compute the gradient at the point .

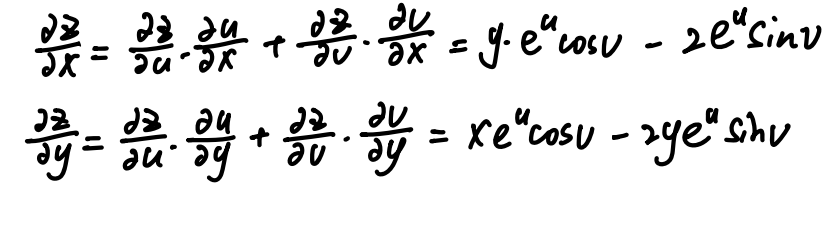


1. What is the geometric relationship between a gradient vector and the function?

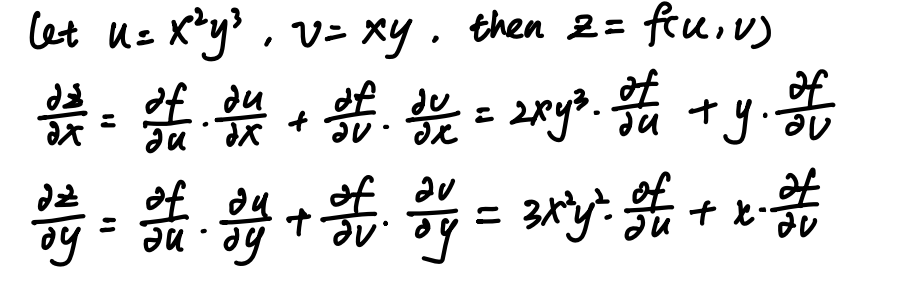
(hint: the gradient is the direction that the … / the gradient is perpendicular to …)



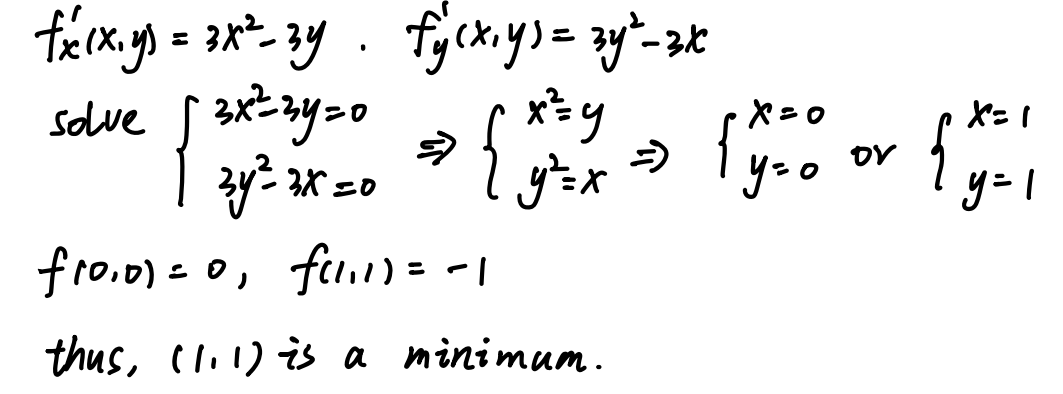
1. Let . Compute , and .



1. Let . Compute , and .



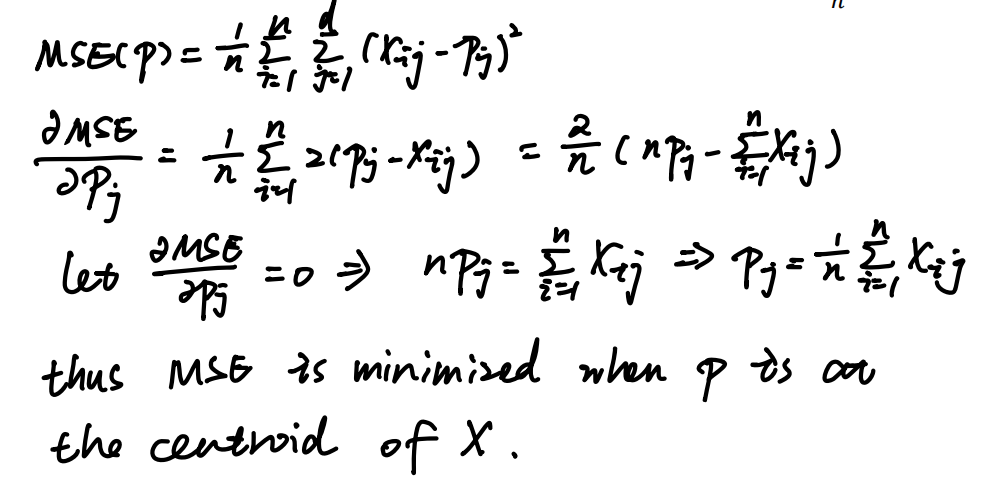
1. Compute the extremum of the function .



1. Suppose there are a set of points and a random point in -dimentional space.

Let the Mean Square Error (MSE): , where is the Euclidean distance between and :

Prove that the is minimized when is at the centroid () of the.



1. Explain the following terms (within 30 words respectively):
   1. supervised learning
   2. unsupervised learning
   3. TP, FP, TN, FN, accuracy, precision, recall and f-score
   4. underfitting and overfitting

(a) supervised learning: Learning from labeled data where correct outputs are known.

(b) unsupervised learning: Learning patterns from data without labels.

(c) TP, FP, TN, FN, Accuracy, Precision, Recall, F1-score:

True/False Positives/Negatives; Accuracy = correct predictions rate; Precision = TP/(TP+FP); Recall = TP/(TP+FN); F1-score = harmonic mean of Precision and Recall.

(d) underfitting and Overfitting:

Underfitting happens when the model is too simple to learn the data patterns.

Overfitting happens when the model is too complex and learns noise along with the data and performs well on training dataset but performs poor on test dataset.

1. Explain the following terms
   1. bias–variance dilemma
   2. No Free Lunch Theorem

(a) It describes the trade-off between a model's ability to fit training data (low bias) and its ability to generalize to unseen data (low variance).

(b) No machine learning algorithm performs best for all problems; algorithm performance depends on the specific problem.

1. What is the difference between validation set and test set?

Validation set is used during training to tune model hyperparameters and prevent overfitting.

Test set is used after training to evaluate the final model’s generalization performance.

1. Please name three main challenges in machine learning.
   1. Overfitting: Overfitting happens when the model is too complex and learns noise along with the data and performs well on training dataset but performs poor on test dataset.
   2. Databias: training data may have hidden biases, leading to unfair or skewed models.
   3. Generalization: models often perform poorly on unseen data even if they fit the training data well.

**Programming part**

1. Get started with a programming language and learn to run matrix operations using your programming language.

(You can submit your programming files. Or you can organize your program into a markdown file or a Jupiter Notebook file as a report and submit it.)

We recommend using the NumPy library in Python, as the TA will use NumPy as an example to explain the programming part. Also, if you plan to use deep learning for research in the future, Python is also a mainstream programming language in academia.

Please pack your files using **zip** and name it in the format of “studentID\_name\_NN\_assignmentID.zip” (e.g. 44278212\_Archimedes\_NN\_1.zip), and

Submit it to TA (Xijian RUI): xijian.rui@fuji.waseda.jp