Recitation 1
Background

10-301/10-601: Introduction to Machine Learning

09/02/2022

1 NumPy and Workflow

NumPy Notebook

Workflow Presentation

Logging Notebook

2 Vectors, Matrices, and Geometry

1. **Inner Product:** $\mathbf{u} = \begin{bmatrix} 6 & 1 & 2 \end{bmatrix}^T, \mathbf{v} = \begin{bmatrix} 3 & -10 & -2 \end{bmatrix}^T$, what is the inner product of $\mathbf{u}$ and $\mathbf{v}$? What is the geometric interpretation?

2. **Cauchy-Schwarz inequality** (Optional): Given $\mathbf{u} = \begin{bmatrix} 3 & 1 & 2 \end{bmatrix}^T, \mathbf{v} = \begin{bmatrix} 3 & -1 & 4 \end{bmatrix}^T$, what is $||\mathbf{u}||_2$ and $||\mathbf{v}||_2$? What is $\mathbf{u} \cdot \mathbf{v}$? How do $\mathbf{u} \cdot \mathbf{v}$ and $||\mathbf{u}||_2 ||\mathbf{v}||_2$ compare? Is this always true?

3. **Matrix algebra.** Generally, if $\mathbf{A} \in \mathbb{R}^{M \times N}$ and $\mathbf{B} \in \mathbb{R}^{N \times P}$, then $\mathbf{AB} \in \mathbb{R}^{M \times P}$ and $(AB)_{ij} = \sum_k A_{ik} B_{kj}$.

   Given $\mathbf{A} = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 4 & -3 & 2 \\ 1 & 1 & -1 \\ 3 & -2 & 2 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

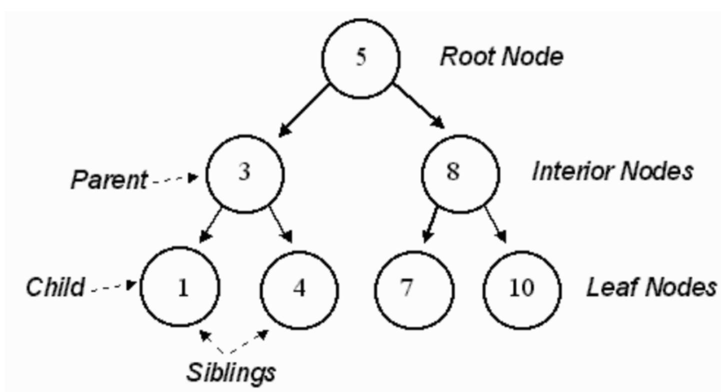
   - What is $\mathbf{AB}$? Does $\mathbf{BA} = \mathbf{AB}$? What is $\mathbf{Bu}$?
   - What is rank of $\mathbf{A}$?
   - What is $\mathbf{A}^T$?
   - Calculate $\mathbf{u}\mathbf{v}^T$.
   - What are the eigenvalues of $\mathbf{A}$?

4. **Geometry:** Given a line $2x + y = 2$ in the two-dimensional plane,
   - If a given point (α, β) satisfies $2\alpha + \beta > 2$, where does it lie relative to the line?
   - What is the relationship of vector $\mathbf{v} = [2, 1]^T$ to this line?
   - What is the distance from origin to this line?

3 CS Fundamentals

1. For each (f, g) functions below, is $f(n) \in \mathcal{O}(g(n))$ or $g(n) \in \mathcal{O}(f(n))$ or both?
   - $f(n) = \log_2(n)$, $g(n) = \log_3(n)$
   - $f(n) = 2^n$, $g(n) = 3^n$
   - $f(n) = \frac{n}{50}$, $g(n) = \log_{10}(n)$
   - $f(n) = n^2$, $g(n) = 2^n$

2. Find the DFS Pre-Order, In-Order, Post-Order and BFS traversal of the following binary tree. What are the time complexities of the traversals?



4 Calculus

1. If $f(x) = x^3 e^x$, find $f'(x)$.

2. If $f(x) = e^x$, $g(x) = 4x^2 + 2$, find $h'(x)$, where $h(x) = f(g(x))$.

3. If $f(x, y) = y \log(1 - x) + (1 - y) \log(x)$, $x \in (0, 1)$, evaluate $\frac{\partial f(x,y)}{\partial x}$ at the point $(\frac{1}{2}, \frac{1}{2})$.

4. Find $\frac{\partial}{\partial w_j} \mathbf{x}^T \mathbf{w}$, where $\mathbf{x}$ and $\mathbf{w}$ are M-dimensional real-valued vectors and $1 \leq j \leq M$.

5 Probability and Statistics

You should be familiar with event notations for probabilities, i.e. $P(A \cup B)$ and $P(A \cap B)$, where A and B are binary events.

In this class, however, we will mainly be dealing with random variable notations, where A and B are random variables that can take on different states, i.e. a_1, a_2, and b_1, b_2, respectively. Below are some notation equivalents, as well as basic probability rules to keep in mind.

- $P(A = a_1 \cap B = b_1) = P(A = a_1, B = b_1) = p(a_1, b_1)$
- $P(A = a_1 \cup B = b_1) = \sum_{b \in B} p(a_1, b) + \sum_{a \in A} p(a, b_1) - p(a_1, b_1)$
- $p(a_1 \mid b_1) = \frac{p(a_1, b_1)}{p(b_1)}$
- $p(a_1) = \sum_{b \in B} p(a_1, b)$

1. Two random variables, A and B, each can take on two values, a_1, a_2, and b_1, b_2, respectively. a_1 and b_2 are considered disjoint (mutually exclusive). $P(A = a_1) = 0.5$, $P(B = b_2) = 0.5$.
   - What is $p(a_1, b_2)$?
   - What is $p(a_1, b_1)$?
   - What is $p(a_1 \mid b_2)$?

2. Now, instead, a_1 and b_2 are not disjoint, but the two random variables A and B are independent.
   - What is $p(a_1, b_2)$?
   - What is $p(a_1, b_1)$?
   - What is $p(a_1 \mid b_2)$?

3. A student is looking at her activity tracker (Fitbit/Apple Watch) data and she notices that she seems to sleep better on days that she exercises. They observe the following:

Exercise	Good Sleep	Probability
yes	yes	0.3
yes	no	0.2
no	no	0.4
no	yes	0.1

   - What is the $P(GoodSleep = yes \mid Exercise = yes)$?
   - Why doesn't $P(GoodSleep = yes, Exercise = yes) = P(GoodSleep = yes) \cdot P(Exercise = yes)$?
   - The student merges her activity tracker data with her food logs and finds that the $P(Eatwell = yes \mid Exercise = yes, GoodSleep = yes)$ is 0.25. What is the probability of all three happening on the same day?

4. What is the expectation of X where X is a single roll of a fair 6-sided dice ($S = \{1, 2, 3, 4, 5, 6\}$)? What is the variance of X?

5. Imagine that we had a new dice where the sides were $S = \{3, 4, 5, 6, 7, 8\}$. How do the expectation and the variance compare to our original dice?