THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

SEAS 8510
Analytical Methods for Machine Learning

Lecture 2

Zachary Dennis, Ph.D.

Agenda

9:00 – 9:20	|	Discussion Groups (20 min)
9:20 – 9:40	|	Homework 1 Review (20 min)
9:40 – 10:35	|	Systems of Linear Equations and Matrix Operations (55 min)
10:35 – 10:45	|	*BREAK (10 min)*
10:45 – 11:45	|	Matrix Forms and Vector Spaces (1 hr)
11:45 – 12:00	|	Homework 2/Discussion 2 Overview (15 min)

Assignments

Last week: Discussion 1/Homework 1

This week: Discussion 2/Homework 2 – Due on 4/6 at 9:00 AM

Discussion 1

Prompt: Describe your past experience and level of knowledge in machine learning, linear algebra, statistics, and Python. How have you used machine learning to help you in a professional role? Comment on how you think improving your knowledge of machine learning will aid you in your career. Provide thoughtful feedback to at least two other students in your discussion group.

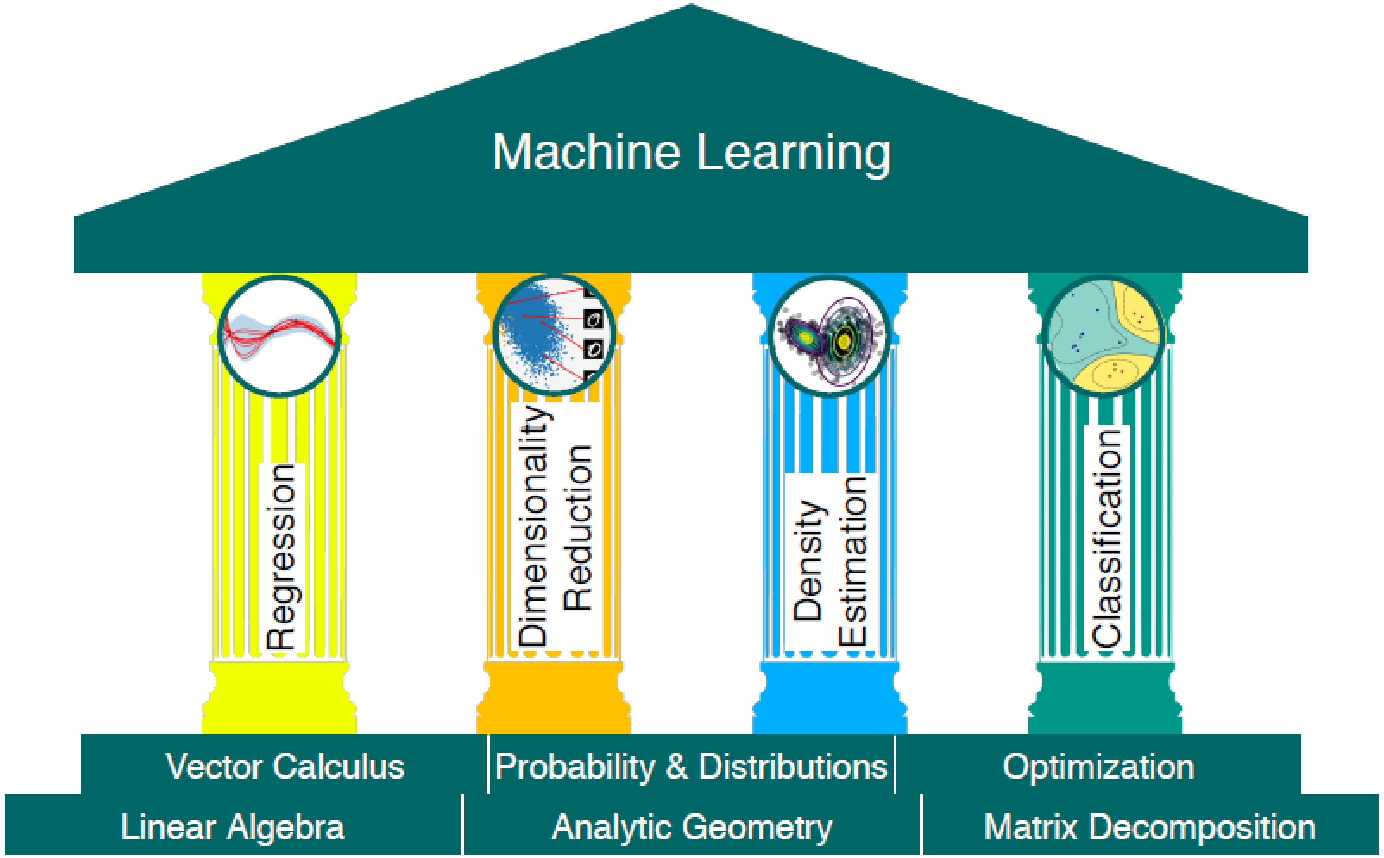
Instructions:

- Choose the breakout room associated with your discussion group number (1-6)
- In small group, discuss your responses (10 min)
- Choose a spokesman to summarize your group's responses (trends, stand out comments, areas of similarity or contrast)
- Back in large group, each spokesman take 90 seconds to summarize your group's responses.

Homework 1

https://colab.research.google.com/drive/1LHpgJRGv3y04FBX8zMBQz6sjCZ5MfO5I?usp=sharing

Mathematical Foundation



- Facilitate creating new machine learning solutions
- Understand and debug existing approaches
- Understand inherent assumptions and limitations of the methodologies

Linear Algebra

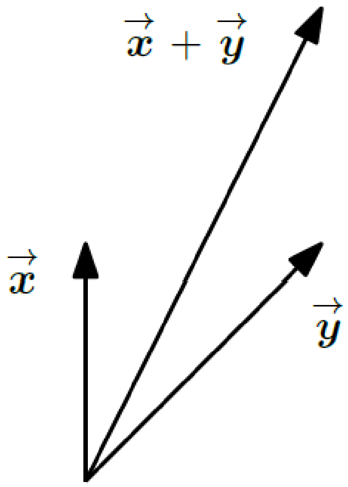
Vectors and Matrices

Vectors

Elements of $\mathbb{R}^n$ (tuples of n real numbers) are vectors.

For a general form vector with n elements, the vector lies in the n-dimensional space $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$



(a) Geometric vectors.

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c} \in \mathbb{R}^n$$

$$\lambda \boldsymbol{a} \in \mathbb{R}^n$$

Systems of Linear Equations

Constraints/System of Linear Equations:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned},$$

$$a_{ij} \in \mathbb{R} \text{ and } b_i \in \mathbb{R}$$

n-tuple unknowns: $$(x_1, \ldots, x_n) \in \mathbb{R}^n$$

The solution is the set of all n-tuples that satisfy the system of linear equations.

Examples!

Matrices and Systems of Linear Equations

Coefficients -> Vectors

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Vectors -> Matrices

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned} \longrightarrow \begin{bmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix}$$

Matrices

Matrix Definition

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}$$

Matrix Addition

$$\boldsymbol{A} + \boldsymbol{B} := \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

Matrix Multiplication

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B} \in \mathbb{R}^{m \times k}$$

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, k$$

Matrix Properties

Identity Matrix

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

$$I_m A = A I_n = A$$

Associative $(AB)C = A(BC)$

$$\lambda(BC) = (\lambda B)C = B(\lambda C) = (BC)\lambda$$

Distributive

$$(A + B)C = AC + BC$$

$$A(C + D) = AC + AD$$

$$\lambda(B + C) = \lambda B + \lambda C$$

Inverse and Transpose

Inverse

$$\boldsymbol{A} * \boldsymbol{A}^{-1} = \boldsymbol{I}$$

Transpose

$$\boldsymbol{B} = \boldsymbol{A}^\top$$

$$b_{ij} = a_{ji}$$

Properties

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} = \boldsymbol{A}^{-1}\boldsymbol{A}$$

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

$$(\boldsymbol{A}+\boldsymbol{B})^{-1} \neq \boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$$

$$(\boldsymbol{A}^\top)^\top = \boldsymbol{A}$$

$$(\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top\boldsymbol{A}^\top$$

$$(\boldsymbol{A}+\boldsymbol{B})^\top = \boldsymbol{A}^\top+\boldsymbol{B}^\top$$

Solving Systems of Linear Equations

Elementary Transformations

1. Exchange of two equations (rows in the matrix representing the system of equations)
2. Multiplication of an equation (row) with a constant $\lambda \in R \setminus \{0\}$
3. Addition of two equations (rows)

Particular Solutions

1. Use elementary transformations to put in Row-Echelon form
2. Set non-pivot column coefficients to 0.

General Solutions

1. Use elementary transformations to put in reduced row-echelon form. AKA Gaussian Elimination
2. Express nonpivot columns as linear combinations of pivot columns

Matrix Forms

Row-Echelon Form

- All rows that contain only zeros are at the bottom of the matrix; correspondingly, all rows that contain at least one nonzero element are on top of rows that contain only zeros.
- Looking at nonzero rows only, the first nonzero number from the left (also called the *pivot* or the *leading coefficient*) is always strictly to the right of the pivot of the row above it.

Use for Specific Solution!

Reduced Row Echelon Form

- It is in row-echelon form.
- Every pivot is 1.
- The pivot is the only nonzero entry in its column.

Use for General Solution!

Calculating Matrix Inverses

$$\boldsymbol{A}^{-1}$$
$$\boldsymbol{A} * \boldsymbol{A}^{-1} = \boldsymbol{I}$$
$$\boldsymbol{A} * \boldsymbol{X} = \boldsymbol{I}$$

$$[\boldsymbol{A}|\boldsymbol{I}_n] \quad \rightsquigarrow \cdots \rightsquigarrow \quad [\boldsymbol{I}_n|\boldsymbol{A}^{-1}]$$

Vector Spaces

$\vec{a} \in \mathbf{V}$: $\vec{a}$ is an element of vector space **V**

All elements of V abide by properties of vectors

$\vec{a} + \vec{b} = \vec{b} + \vec{a}$

$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$

$\vec{a} + 0 = \vec{a}$

$\vec{a} + -\vec{a} – 0$

$c(\vec{a} + \vec{b}) = c\,\vec{a} + c\,\vec{b}$

$(c + d)\,\vec{a} = c\,\vec{a} + d\,\vec{a}$

$cd(\vec{a}) = c(d\,\vec{a})$

$1\,\vec{a} = \vec{a}$

Vector Spaces: Closure

Vector Space **V** is a collection of elements that can be

1. added together in any combination
2. multiplied by scalars in any combination

Closure:

Given $\vec{\boldsymbol{a}} \in \boldsymbol{V}$ and scalar **c**, then $\mathbf{c}\vec{\boldsymbol{a}} \in \boldsymbol{V}$

Given $\vec{\boldsymbol{a}} \in \boldsymbol{V}$ and $\vec{\boldsymbol{b}} \in \boldsymbol{V}$, then $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} \in \boldsymbol{V}$

Examples: R, R^n, $R^{m \times n}$

Vector Subspace

Vector Subspace notation: V = (V,+, ▪)

U is vector subspace of V: U $\subseteq$ V, U $\neq \emptyset$. Then U = (U,+, ▪)

U has to contain 0 and be closed to be a subspace

https://textbooks.math.gatech.edu/ila/subspaces.html

Matrices in NumPy

```
A = np.arange(60).reshape(6,10)
sub = A[1:4:1,0:5:1]

D = np.diag(sub)
UT = np.triu(sub)
LT = np.tril(sub)

A = np.array([[1,3],[5,8]])
B = np.array([[3, 2],[1, 8]])
print(A+B)
print(A*B)
print(A@B)
print(np.dot(A,B))

print(np.diag([1,2,3,4,5]))
```

Matrices in NumPy

```
matrix_array = np.array([[1, 2], [3, 4]])
zero_matrix = np.zeros((2, 2))
ones_matrix = np.ones((2, 2))
eye_matrix = np.eye(2)
identity_matrix = np.identity(2)
full_matrix = np.full((2, 2), 7)
random_matrix = np.random.rand(2, 2)
randomn_matrix = np.random.randn(2, 2)
reshaped_matrix = np.reshape(matrix_array, (4, 1))
transposed_matrix = np.transpose(matrix_array)
transposed_matrix_short = matrix_array.T
stacked_v_matrix = np.vstack((matrix_array, ones_matrix))
stacked_h_matrix = np.hstack((matrix_array, ones_matrix))
split_matrices = np.split(matrix_array, 2)
split_h_matrices = np.hsplit(matrix_array, 2)
split_v_matrices = np.vsplit(matrix_array, 2)
determinant = np.linalg.det(matrix_array)
matrix_trace = np.trace(matrix_array)
```