SCHOOL OF COMPUTER TECHNOLOGY

AASD 4001
Mathematical Concepts for Machine Learning
Lecture 1

Reza Moslemi, Ph.D., P.Eng.
reza.moslemi@georgebrown.ca

Session 1

- Introduction to NumPy
- Vector Space Analysis and Linear Algebra

Introduction to Numpy

NumPy (Numerical Python):

- An open source Python library used in almost every field of science and engineering
- A universal standard for working with numerical data in Python
- NumPy users include everyone from beginner coders to experienced researchers doing state-of-the-art scientific and industrial research and development
- NumPy arrays are used extensively in Pandas, SciPy, Matplotlib, scikit-learn, scikit-image and most other data science, and machine learning Python packages.

Introduction to Numpy (cont'd)

The NumPy library contains multidimensional array and matrix data structures

- It provides ndarray, a homogeneous n-dimensional array object, with methods to efficiently operate on it.
- NumPy can be used to perform a wide variety of mathematical operations on arrays.
- It adds powerful data structures to Python that guarantee efficient calculations with arrays and matrices and it supplies an enormous library of high-level mathematical functions that operate on these arrays and matrices.

What is an array?

- An array is a central data structure of the NumPy library
- A grid of values containing information about the raw data, how to locate an element, and how to interpret an element
- The elements are all of the same type, referred to as the array dtype

Introduction to Numpy (cont'd)

Installing NumPy

- conda install numpy
- pip install numpy

Importing NumPy

- import numpy as np

Let's switch to the jupyter notebook and open "Practice 1.ipynb" to practice NumPy!

Vector Space Analysis and Linear Algebra

Linear algebra

- Is the branch of mathematics concerning linear equations. In the context of machine learning, it is the mathematical toolset to work with data (often vectors, matrices, or tensors)

Vector space:

- A vector space (also called a linear space) is a set of vectors, which may be added together and multiplied ("scaled") by scalar numbers.

What are scalars, vectors, matrices and tensors?

Scalar: a scalar value is simply a number

- e.g.: 0.1, -5, 48.2, pi

Vector: an n-by-1 entity with n values (1D)

- e.g.: $\begin{bmatrix} 2 \\ 0 \\ -8.3 \end{bmatrix}$ is a 3-by-1 vector (3 rows and 1 column)

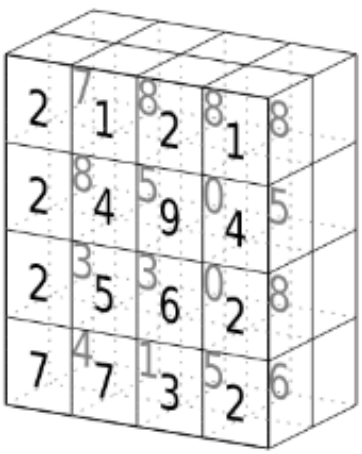
Vector Space Analysis and Linear Algebra

Matrix: an n-by-m entity with n*m values (2D)

- e.g.: $\begin{bmatrix} 4 & 6 & 75 & 8.4 \\ -8 & 5 & 6 & 55 \\ 0 & 0 & 42 & 54 \end{bmatrix}$ is a 3-by-4 matrix (3 rows and 4 column)

Tensor: an n-by-m-by-l entity with n*m*l values (3D)

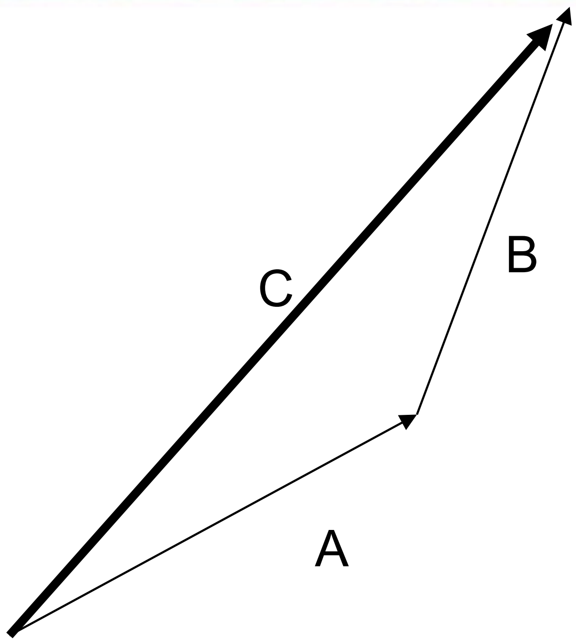
3	1	4	1
5	9	2	6
5	3	5	8
9	7	9	3
2	3	8	4
6	2	6	4



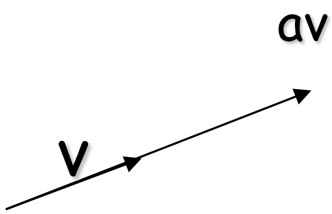
Vector Space Analysis and Linear Algebra

Vector operations:

- Addition: $\begin{bmatrix} 8 \\ 5 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \\ 14 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ 18 \end{bmatrix}$



- Scalar product: $3*\begin{bmatrix} 8 \\ 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 24 \\ 15 \\ 12 \end{bmatrix}$



Vector Space Analysis and Linear Algebra

Vector operations:

- Dot product: $\begin{bmatrix} 8 \\ 5 \\ 4 \end{bmatrix} . \begin{bmatrix} 0 \\ -2 \\ 14 \end{bmatrix} = 8 * 0 + 5 * (-2) + 4 * 14 = 46$
  - The result of vector dot product (aka inner product) is a scalar (just a number!)
  - The dot product of a vector with itself is the square of its magnitude: $\begin{bmatrix} 8 \\ 5 \\ 4 \end{bmatrix} . \begin{bmatrix} 8 \\ 5 \\ 4 \end{bmatrix} =$ $8 * 8 + 5 * 5 + 4 * 4 = 105$
  - The dot product is also related to the angle between the two vectors: $\vec{A}.\vec{B} = |\vec{A}||\vec{B}| \cos\theta$ where θ is the angle btw the two vectors.
  - If $\vec{A}.\vec{B}$ =0, it means that the 2 vectors are perpendicular to each other.

Vector Space Analysis and Linear Algebra

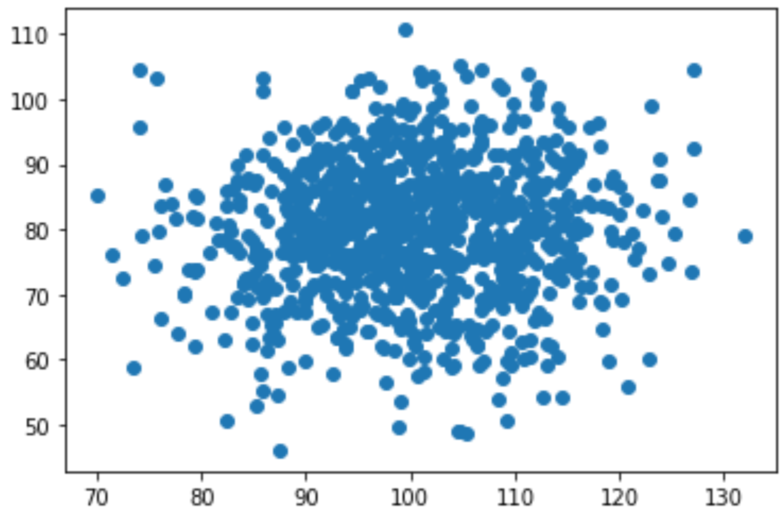
Vector operations:

- Cross product: $\begin{bmatrix}8\\5\\4\end{bmatrix} \times \begin{bmatrix}0\\-2\\14\end{bmatrix} = \begin{bmatrix}5*14-(-2)*4\\-(8*14-4*0)\\8*(-2)-0*5\end{bmatrix} = \begin{bmatrix}78\\-112\\-16\end{bmatrix}$
  - The result of vector cross product is another vector (**NOT** just a number!)
  - The resultant vector is perpendicular to both original vectors.
  - $\begin{bmatrix}8\\5\\4\end{bmatrix}.\begin{bmatrix}78\\-112\\-16\end{bmatrix}$=0 and $\begin{bmatrix}0\\-2\\14\end{bmatrix}.\begin{bmatrix}78\\-112\\-16\end{bmatrix}$=0
  - The cross product is also related to the angle between the two vectors: $\vec{A}\times\vec{B} = |\vec{A}||\vec{B}|\sin\theta\,\vec{n}$ where θ is the angle btw the two vectors and $\vec{n}$ is a unit vector perpendicular to the plane containing $\vec{A}$ and $\vec{B}$.
  - If $\vec{A}\times\vec{B}$ =0, it means that the 2 vectors are collinear, either in the same direction or exact opposite direction.

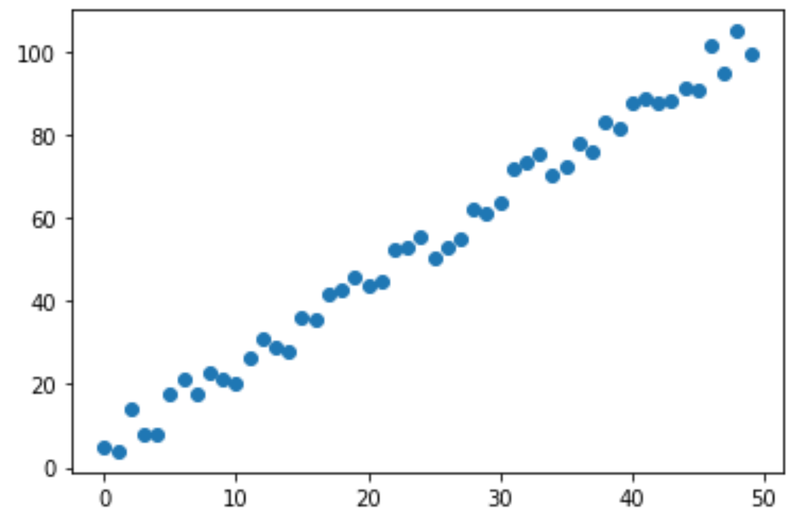
Vector Space Analysis and Linear Algebra

Covariance and correlation

- Covariance is a mathematical term to quantitatively measure how much two vectors are related to each other. Similarly, it measures how two vectors change with respect to one another
  - $cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N-1}$, no need to remember this formula as we will use python to calculate covariance.
  - What is $cov_{x,x}$?



low covariance value



high covariance value

Vector Space Analysis and Linear Algebra

Covariance and correlation

- A problem with covariance is that it is difficult to interpret. What is a large vs. low covariance value? Is 10, 50, 5000 high or low?
- In order to solve that problem, we need a normalized metric, i.e. correlation.

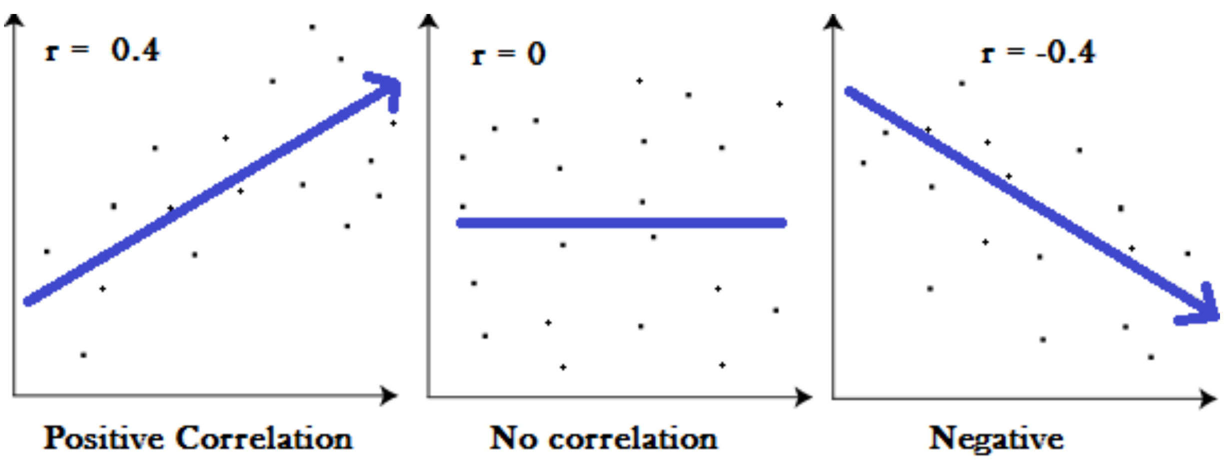
Correlation

- Correlation is obtained from dividing the covariance by the std of both variables
- $\rho_{x,y} = \frac{cov_{x,y}}{\sigma_x \sigma_y}$
- Correlation is always btw -1 and 1.

Vector Space Analysis and Linear Algebra

Correlation

- $\rho_{x,y} = 1$ means that there is perfect correlation.
- $\rho_{x,y} = 0$ means that there is no correlation.
- $\rho_{x,y} = -1$ means that there is perfect inverse correlation.



Vector Space Analysis and Linear Algebra

Matrix operations:

- Element-wise addition: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \pm \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a \pm e & b \pm f \\ c \pm g & d \pm h \end{bmatrix}$
- Multiplication: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} . \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$
- Is AB=BA? Try with A= $\begin{bmatrix} 1 & 2 \\ 7 & 4 \end{bmatrix}$ and B= $\begin{bmatrix} 4 & 1 \\ -2 & 0 \end{bmatrix}$
- $\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$ can be used to scale a vector by α: $\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha x \\ \alpha y \end{bmatrix} = \alpha \begin{bmatrix} x \\ y \end{bmatrix}$

Vector Space Analysis and Linear Algebra

Matrix operations:

- Identity matrix: $I_n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix}$

  1 on the main diagonal and zero elsewhere
- Inverse of a matrix (A^{-1}): $AA^{-1} = A^{-1}A = I_n$
  - The above simple inverse is only defined for square matrices.
  - Even for square matrices, an inverse may NOT always exist.
- Transpose of a matrix (A' or A^T): flipping a matrix over its diagonal; switching the row and column indices of the matrix.
- Determinant of a 2-by-2 matrix: $\det(\mathrm{A}) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} =$
- Determinant of a 3-by-3 matrix: $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} =$

Vector Space Analysis and Linear Algebra

Matrix operations:

- Inverse of a 2-by-2 matrix: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
- Inverse of a 3-by-3 matrix:

$$A^{-1} = \frac{1}{\det(A)}\,\mathrm{adj}(\mathrm{A})$$
$$\mathrm{adj}(\mathrm{A}) = C^T$$
$$C = (-1)^{i+j} M_{ij}$$

Adjugate of a matrix is the transpose of the cofactor matrix.

M_{ij}, the (i, j) minor, is the determinant of the submatrix formed by deleting the i^{th} row and j^{th} column.

What is the inverse of the following matrix?

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} 0.5 & -0.25 & 0 \\ 0 & 0.5 & 0 \\ -1 & 0.5 & 1 \end{bmatrix}$$

Let's switch to jupyter notebook and open "Practice 2.ipynb" to practice what we have learnt!