Fundamentals of Information Systems

Python Programming (for Data Science)

Master's Degree in Data Science

Gabriele Tolomei

gtolomei@math.unipd.it
University of Padua, Italy
2018/2019
November, 12 2018

Lecture 6 (Extra): Basics of Linear Algebra

What is a Matrix?

- A bidimensional array which is the building block of linear algebra.
- Linear algebra is used quite a bit in advanced statistics, largely because it provides two benefits:
    - Compact notation for describing sets of data and sets of equations;
    - Efficient methods for manipulating sets of data and solving sets of equations.

Matrix Definition

- A **matrix** is a rectangular array of numbers arranged in **rows** and **columns**.
- The following is an example of a 3-by-4 matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1.2 & -0.7 & 3.1 & 2.8 \\ -5.9 & 1.4 & 0.3 & -4.3 \\ 0.0 & 1.0 & 12.7 & 6.5 \end{bmatrix}$$

Matrix Definition

- More generally, an m-by-n matrix $\mathbf{A}$ can be represented as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

- a_{ij} refers to the element of $\mathbf{A}$ located at the i-th row and j-th column.
- m and n are called **dimensions** of the matrix.
- Sometimes, you specifiy dimensions when defining a matrix, e.g., $\mathbf{A}_{m,n}$.

Matrix Equality

- Two matrices $\mathbf{A}$ and $\mathbf{B}$ are equal if **all three** of the following conditions are met:
    - Each matrix has the same number of rows;
    - Each matrix has the same number of columns;
    - Corresponding elements within each matrix are equal.

Transpose Matrix

- The transpose of a matrix $\mathbf{A}_{m,n}$ is another matrix $\mathbf{A}^T_{n,m}$ that is obtained by using rows from the first matrix as columns in the second matrix.
- For example, it is easy to see that the transpose of matrix $\mathbf{A}_{3,2}$ is $\mathbf{A}^T_{2,3}$:

$$\mathbf{A} = \begin{bmatrix} 1.2 & -0.7 \\ -5.9 & 1.4 \\ 0.0 & 1.0 \end{bmatrix} \quad \mathbf{A}^T = \begin{bmatrix} 1.2 & -5.9 & 0.0 \\ -0.7 & 1.4 & 1.0 \end{bmatrix}$$

- Row 1 of matrix $\mathbf{A}$ becomes column 1 of $\mathbf{A}^T$, row 2 of $\mathbf{A}$ becomes column $\mathbf{A}^T$, and finally row 3 of $\mathbf{A}$ becomes column 3 of $\mathbf{A}^T$.

Vectors

- Vectors are a "special" type of matrix, which have only one column or one row.
- They come in **two** flavors: **column vectors** and **row vectors**.
- For example, matrix $\mathbf{a}$ is a 3-by-1 column vector, and matrix $\mathbf{a}^T$ is a 1-by-3 row vector.

$$\mathbf{a} = \begin{bmatrix} 1.2 \\ -5.9 \\ 0.0 \end{bmatrix} \quad \mathbf{a}^T = \begin{bmatrix} 1.2 & -5.9 & 0.0 \end{bmatrix}$$

Square Matrix

- A square matrix is a matrix having the same number of rows and columns (i.e., an n-by-n matrix).
- Some kinds of square matrices are particularly interesting:
    - **Symmetric Matrix**
    - **Diagonal Matrix**
    - **Scalar Matrix**

Symmetric Matrix

- A matrix $\mathbf{A}_{n,n}$ is **symmetric** if its transpose $\mathbf{A}_{n,n}^T$ is equal to itself.
- For example:

$$\mathbf{A} = \begin{bmatrix} 1.2 & -5.9 \\ -5.9 & 1.2 \end{bmatrix} = \begin{bmatrix} 1.2 & -5.9 \\ -5.9 & 1.2 \end{bmatrix} = \mathbf{A}^T$$

Diagonal Matrix

- A **diagonal** matrix $\mathbf{A}_{n,n}$ is a special type of **symmetric** matrix, in which it has zeros in the off-diagonal elements.
- For example:

$$\mathbf{A} = \begin{bmatrix} 1.2 & 0 & 0 \\ 0 & 2.7 & 0 \\ 0 & 0 & -3.1 \end{bmatrix}$$

Scalar Matrix

- A **scalar** matrix $\mathbf{A}_{n,n}$ is a special kind of **diagonal** matrix, in which it has equal-valued elements along the diagonal.
- For example:

$$\mathbf{A} = \begin{bmatrix} 2.7 & 0 & 0 \\ 0 & 2.7 & 0 \\ 0 & 0 & 2.7 \end{bmatrix}$$

Matrix Operations

Matrix Addition and Subtraction

- Just like ordinary algebra, linear algebra has operations like addition and subtraction.
- Two matrices can be added or subtracted **only if** they have the same dimensions, i.e., the same number of rows and columns.
- Addition or subtraction is accomplished **element-wise**. For example, consider the following matrices $\mathbf{A}$ and $\mathbf{B}$.

$$\mathbf{A} = \begin{bmatrix} 1.2 & -0.7 & 9.8 \\ -5.9 & 1.4 & 6.2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -0.8 & -2.9 & 0.0 \\ 1.6 & 1.4 & 1.0 \end{bmatrix}$$

Matrix Addition and Subtraction

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 0.4 & -3.6 & 9.8 \\ -4.3 & 2.8 & 7.2 \end{bmatrix} \qquad \mathbf{A} - \mathbf{B} = \begin{bmatrix} 2.0 & 2.2 & 9.8 \\ -7.5 & 0.0 & 5.2 \end{bmatrix}$$

- Note that addition is commutative (i.e., $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$), but subtraction in general is not.

Matrix Multiplication

- In linear algebra, there are **two** kinds of matrix multiplication:
    - multiplication of a matrix by a scalar (i.e., a number);
    - multiplication of a matrix by another matrix.

How to Multiply a Matrix by a Scalar

- When you multiply a matrix $\mathbf{A}$ by a scalar, you multiply **every element** in the matrix by that same number.
- This operation produces a new matrix, which is called a **scalar multiple**.
- For example, consider the following:

$$\mathbf{A} = \begin{bmatrix} 1 & 9 & 4 \\ 5 & 2 & 0 \\ -1 & 3 & 3 \end{bmatrix} \quad k \cdot \mathbf{A} = \begin{bmatrix} k & 9k & 4k \\ 5k & 2k & 0 \\ -1k & 3k & 3k \end{bmatrix} \quad (k \in \mathbb{R})$$

How to Multiply a Matrix by a Matrix

- The product of a matrix $\mathbf{A}$ by another matrix $\mathbf{B}$, i.e., $\mathbf{A} \cdot \mathbf{B}$ is defined **only** when the number of columns in $\mathbf{A}$ is equal to the number of rows in $\mathbf{B}$.
- Analogously, $\mathbf{B} \cdot \mathbf{A}$ is defined only when the number of columns in $\mathbf{B}$ is equal to the number of rows in $\mathbf{A}$.
- More generally, if $\mathbf{A}$ is an m-by-k matrix, and $\mathbf{B}$ is an k-by-n matrix the matrix product $\mathbf{A} \cdot \mathbf{B}$ is an m-by-n matrix $\mathbf{C}$.
- Each element of $\mathbf{C}$ can be therefore computed according to the following formula:

$$c_{ij} = \sum_{p=1}^{k} a_{ip} \cdot b_{pj}$$

How to Multiply a Matrix by a Matrix

- In the formula above we identify:
  - c_{ij} as the element in row i and column j of the resulting matrix $\mathbf{C}$;
  - a_{ip} as the element in row i and column p of the first operand matrix $\mathbf{A}$;
  - b_{pj} as the element in row p and column j of the second operand matrix $\mathbf{B}$;
  - $\sum_{p=1}^{k}$ indicates that $a_{ip} \cdot b_{pj}$ must be summed over $p = 1 \ldots k$.

Matrix Multiplication: An Example

- Let's work through an example to show how the above formula works. Suppose we want to compute $\mathbf{A} \cdot \mathbf{B}$, given the matrices below:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 6 & 7 \\ 8 & 9 \\ 10 & 11 \end{bmatrix}$$

- Let $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$, which we know will be a 2-by-2 matrix.

Matrix Multiplication: An Example

$$c_{11} = \sum_{p=1}^{3} a_{1p} \cdot b_{p1} = 0 * 6 + 1 * 8 + 2 * 10 = 0 + 8 + 20 = 28$$

$$c_{12} = \sum_{p=1}^{3} a_{1p} \cdot b_{p2} = 0 * 7 + 1 * 9 + 2 * 11 = 0 + 9 + 22 = 31$$

$$c_{21} = \sum_{p=1}^{3} a_{2p} \cdot b_{p1} = 3 * 6 + 4 * 8 + 5 * 10 == 18 + 32 + 50 = 100$$

$$c_{22} = \sum_{p=1}^{3} a_{2p} \cdot b_{p2} = 3 * 7 + 4 * 9 + 5 * 11 = 21 + 36 + 55 = 112$$

Matrix Multiplication: An Example

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{C} = \begin{bmatrix} 28 & 31 \\ 100 & 112 \end{bmatrix}$$

Multiplication Order

- In some cases, matrix multiplication is defined for $\mathbf{A} \cdot \mathbf{B}$, but not for $\mathbf{B} \cdot \mathbf{A}$, and vice versa.
- However, even when matrix multiplication is possible in both directions, results may be different. That is, $\mathbf{A} \cdot \mathbf{B}$ is generally different from $\mathbf{B} \cdot \mathbf{A}$.
- The bottom line: when you multiply two matrices, order matters!

Identity Matrix

- The **identity matrix** is an n-by-n diagonal matrix with 1's in the diagonal and 0's everywhere else.
- The identity matrix is often denoted by $\mathbf{I}$ (or $\mathbf{I}_{n,n}$ or $\mathbf{I}_n$).
- The identity matrix has a nice property: Any matrix that can be multiplied by $\mathbf{I}$ remains the same, that is:

$$\mathbf{A} \cdot \mathbf{I} = \mathbf{I} \cdot \mathbf{A} = \mathbf{A}$$

- Of course, if $\mathbf{A}$ is not a square matrix, $\mathbf{I}$ will have different size depending on whether you do $\mathbf{A} \cdot \mathbf{I}$ or $\mathbf{I} \cdot \mathbf{A}$.

Vector Multiplication

- The multiplication of a vector by a vector produces some interesting results.
- One is known as the vector **inner product** (a.k.a. **dot product** or **scalar product**), whilst the other is called the vector **outer product**.

Vector Inner Product (Dot Product)

- Assume that $\mathbf{a}$ and $\mathbf{b}$ are vectors, each with the same number of elements n. Then, the **inner product** of $\mathbf{a} \cdot \mathbf{b}$ is a scalar $s \in \mathbb{R}$.
$$\mathbf{a}^T \cdot \mathbf{b} = \mathbf{b}^T \cdot \mathbf{a} = s$$
- $\mathbf{a}$ and $\mathbf{b}$ are column vectors, each having n elements;
- $\mathbf{a}^T$ is the transpose of $\mathbf{a}$, which makes $\mathbf{a}^T$ a row vector;
- $\mathbf{b}^T$ is the transpose of $\mathbf{b}$, which makes $\mathbf{b}^T$ a row vector;
- s is a scalar; that is, s is a real number, **not** a matrix!
- Note that the product of two matrices is usually another matrix. However, the inner product of two vectors is a real number!

Vector Outer Product

- Assume that $\mathbf{a}$ and $\mathbf{b}$ are vectors of m and n elements, respectively. Then, the **outer product** of $\mathbf{a} \otimes \mathbf{b}$ is an m-by-n matrix $\mathbf{C}$.

$$\mathbf{a} \otimes \mathbf{b}^T = \mathbf{C}$$

- $\mathbf{a}$ is an m-by-1 column vector;
- $\mathbf{b}^T$ is the transpose of $\mathbf{b}$, which makes $\mathbf{b}^T$ a 1-by-n row vector;
- $\mathbf{C}$ is an m-by-n matrix.
- Let's see how this works!

Vector Outer Product

$$\mathbf{a} = \begin{bmatrix} u \\ v \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \mathbf{a} \otimes \mathbf{b}^T = \mathbf{C} = \begin{bmatrix} u \cdot x & u \cdot y & u \cdot z \\ v \cdot x & v \cdot y & v \cdot z \end{bmatrix}$$

- Notice that the elements of matrix $\mathbf{C}$ consist of the product of elements from vector $\mathbf{a}$ "crossed" with elements from vector $\mathbf{b}$.

Norm of a Vector

- A **norm** is a function that assigns a strictly positive length to a vector (in a vector space).
- Given a vector $\mathbf{x} \in \mathbb{R}^n = (x_1, \ldots, x_n)$ we define the ℓ_p-norm (a.k.a. the p-norm), with $p \geq 1$ as follows:

$$||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

where $|x_i|$ is the **absolute value** of x_i, and $|x_i| = x_i$ iff $x_i \geq 0$; $-x_i$, otherwise.

ℓ_p-norm

- ℓ_1 $(p = 1)$ a.k.a. the **taxicab norm** or **Manhattan norm**:

$$||\mathbf{x}||_1 = |\mathbf{x}| = \sum_{i=1}^{n} |x_i|$$

- ℓ_2 $(p = 2)$ a.k.a. the **Euclidean norm**:

$$||\mathbf{x}||_2 = ||\mathbf{x}|| = \sqrt{x_1^2 + \ldots + x_n^2}$$

- ℓ_∞ $(p = \infty)$ as p approaches to ∞ the p-norm approaches the **infinity norm** or **maximum norm**:

$$||\mathbf{x}||_\infty = \max_i |x_i|$$