Graph Theory, Discrete Mathematics and Optimization – *Module Graph Theory and Discrete Mathematics*

Giovanni Naldi — act I

Alessandra Micheletti — act II

Roberto Cordone — act III

Main Web page: **Ariel site**

Exam: 2 mid term (after I and II part + after III part) OR written exam with I+II+III part

Module Graph Theory and Discrete Mathematics – I part

Topics:

- ❑ (some) Linear Algebra
- ❑ Introduction to Graph Theory
- ❑ Difference equations and dynamical systems

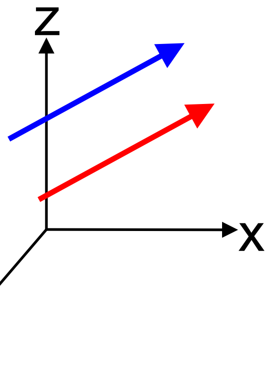
Linear algebra is an area of study in Mathematics that concerns itself primarily with the study of vector spaces and the linear transformations between them.

What? And the data?

wait wait … Linear algebra is behind (almost) all the powerful machine learning algorithms and vectors/matrices/tensors are a basic data structure

Scalars, Vectors, and Tensors

- A Scalar
  - Has magnitude only (**e.g. T=temperature**)
  - Represented by a single number
- A Scalar <u>Field</u>
  - A scalar as function of position (e.g. T=T(x,y,z))
  - Represented by a single number whose value varies in space.
- A Vector
  - Characterized by a magnitude and direction (**e.g. $\mathbf{v}$=velocity**)
  - Represented by a set of numbers (e.g. in 3 dimensions 3 numbers)
  - Represented as an arrow with length and spatial orientation
  - Two vectors are said to be equal if they are Parallel (Pointed in same direction) and of equal length (magnitude).
- A Vector <u>Field</u>
  - A vector whose magnitude and direction vary in space (e.g. $\mathbf{v}=\mathbf{v}(x,y,z)$).



Two Equal Vectors

Scalars, Vectors, and Tensors

- A tensor (here we refer **only** to three-dimensional space)
  - Characterized by an *order*.
  - In general then:
    - Zeroth-order tensor is a scalar
    - First-order tensor is a vector
    - Second order tensor looks like a 3x3 matrix.
  - An n^{th} order tensor has 3^n components
  - Usually, “tensor” refers to a second order tensor
    - Ordered set of nine numbers, each of which is associated with two directions
    - “Arrow-in-space” concept not helpful
    - Stress tensor a common example in fluid mechanics

Abstract version

A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below.[1] The axioms must hold for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and for all scalars c and d.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a **zero** vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$, is in V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

Example. scalar: real numbers; vector: a force.

For this course we will consider only space the Euclidean Vector Spaces

Note. Symbol $\mathbb{R}$ for the set of the real numbers.

Vectors in the Euclidean n-space, usually denoted by $\mathbb{R}^n$, are the ordered n-tuples: sequence of n real numbers. The standard notation for a vector is a boldface lowercase letter and the n-tuple is represented as a column vector, as in

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

In general, scalar multiplication and addition in $\mathbb{R}^n$. are, respectively, defined by

$$\alpha\mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} \quad \text{and} \quad \mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

for any **x**, **y** ∈ $\mathbb{R}^n$. and any scalar α.

Example if $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$, then

$$-x = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \quad 3x = \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \quad -2x = \begin{bmatrix} -4 \\ -2 \end{bmatrix}$$

$$\mathbf{x}+\mathbf{y} = \begin{pmatrix} 5 \\ -3 \end{pmatrix} \qquad \mathbf{x}\text{-}\mathbf{y} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$$

Example consider the following values for a Iris flower:

- sepal length
- sepal width
- petal length
- petal width

For the iris-setosa we have the vector:

$$\mathbf{is} = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}$$

The term *matrix* means simply a rectangular array of numbers. A matrix having *m* rows and *n* columns is said to be *m*×*n*. A matrix is said to be *square* if it has the same number of rows and columns, that is, if *m* = *n*.

Notation. The name of a matrix is usually a capital letter, the elements (entries) a_{ij} have two indices: i row number, j column number,

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

We say that two matrices are **equal** if they have the same size (i.e., the same number of rows and the same number of columns) and if their corresponding entries are equal. If A and B are mxn matrices, then the **sum** A =C+B is the mxn matrix whose entries are the sum of the corresponding entries in A and B. The sum is defined only when A and B are the same size.

If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose entries are r times the corresponding entries in A.

Example

L. Eldén. Numerical linear algebra in data mining. *Acta Numerica*, 2006,

Example 1.1 [1] *Term-document matrices* are used in *information retrieval*. Consider the following selection of four documents. Key words, which we call *terms*, are marked in boldface[2].

Document 1: The **Google matrix** P is a model of the **Internet**.
Document 2: P_{ij} is nonzero if there is a **link** from **web page** j to i.
Document 3: The **Google matrix** is used to **rank** all **web pages**
Document 4: The **ranking** is done by solving a **matrix eigenvalue** problem. If we
Document 5: **England** dropped out of the top 10 in the **FIFA** **ranking**.

count the frequency of terms in each document we get the following result.

Term	Doc 1	Doc 2	Doc 3	Doc 4	Doc 5
eigenvalue	0	0	0	1	0
England	0	0	0	0	1
FIFA	0	0	0	0	1
Google	1	0	1	0	0
Internet	1	0	0	0	0
link	0	1	0	0	0
matrix	1	0	1	1	0
page	0	1	1	0	0
rank	0	0	1	1	1
web	0	1	1	0	0

Thus each each document is represented by a vector, or a point, in $\mathbb{R}^8$, and we can organize them as a term-document matrix,

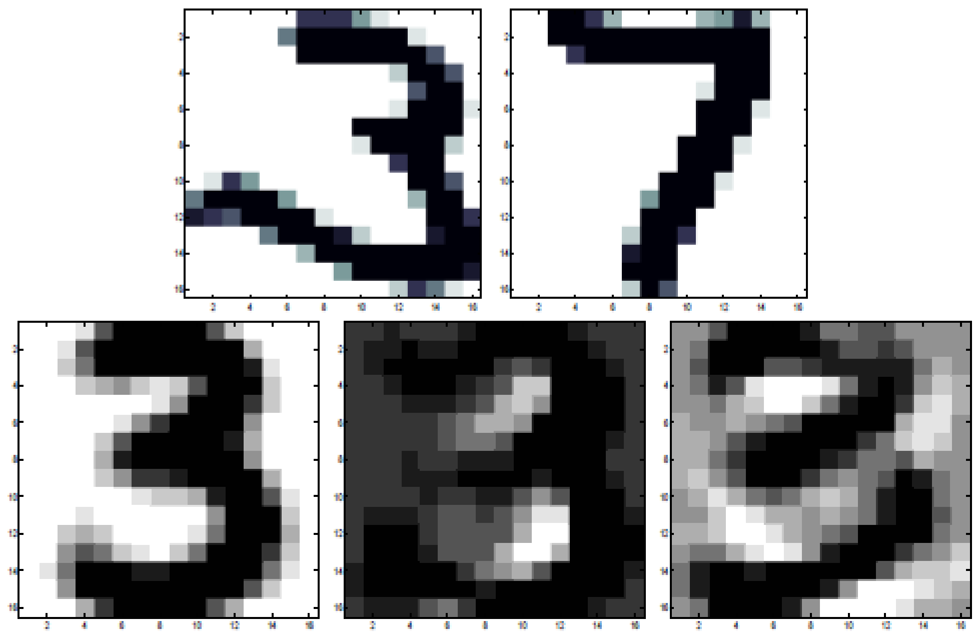
$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Now assume that we want to find all documents that are relevant to the query "**ranking of web pages**". This is represented by *query* vector, constructed in an analogous way as the term-document matrix,

$$q = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^{10}.$$

Thus the query itself is considered as a document. The information retrieval task can now be formulated as a matematical problem: *find the columns of A that are close to the vector q*. To solve this problem we must use some distance measure in $\mathbb{R}^{10}$.

Example 1.2 In handwritten digit recognition vectors are used to represent digits. The image of one digit is a 16×16 matrix of numbers, representing grey scale. It can also be represented as a vector in $\mathbb{R}^{256}$, by stacking the columns of the matrix. A set of n digits (handwritten 3's, say) can then be represented by matrix $A \in \mathbb{R}^{256 \times n}$, and the columns of A span a subspace of $\mathbb{R}^{256}$.

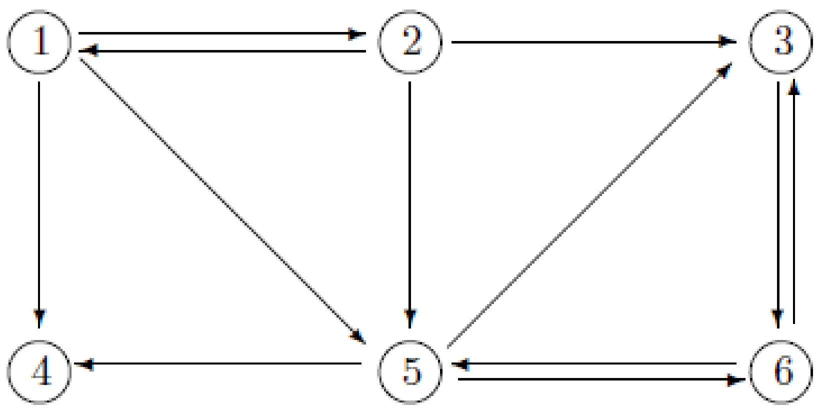


Handwritten digits from the US Postal Service data base

Example 1.3 The task of extracting information from all the web pages available on the Internet, is done by *search engines*. The core of the Google search engine is a matrix computation, probably the largest that is performed routinely

The matrix is constructed based on the link structure of the web,

The following small link graph illustrates a set of web pages with outlinks and inlinks.

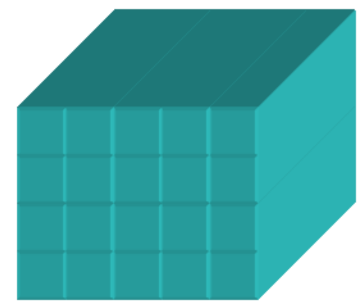


The corresponding matrix becomes

$$P = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

Tensor: “Generalization of an n-dimensional array”

Vector: order-1 tensor



Order-3 tensor

Matrix: order-2 tensor