

## 2

### Algebra lineare di base

Con il termine data science si intende un insieme di metodi statistici per lo studio di una pluralità di variabili quantitative e/o qualitative rilevate su più unità. L'obiettivo dell'analisi dei dati è la ricerca di una struttura "pattern" nei dati che, in prima approssimazione, può essere costituita dall'accertamento delle relazioni tra le variabili e dall'individuazione di somiglianze tra le unità. Il punto di partenza per l'analisi è la tabella (matrice)  $X$  di dimensione  $n \times p$  dove  $x_{is}$  (l'elemento che è all'incrocio della riga  $i$  e della colonna  $s$  indica la modalità o codice nella  $i$ -esima unità statistica della variabile  $s$ -esima ( $i = 1, \dots, n$ ;  $s = 1, 2, \dots, p$ ). Ad esempio nel caso del file di input `Firm.xlsx` la matrice dei dati di partenza è di dimensione  $107 \times 9$  (la prima colonna "Code" contiene i nomi delle righe). E' necessario, quindi, introdurre una serie di operazioni preliminari sulle matrici.

#### 2.1 Operazioni elementari con le matrici

Una matrice  $A$  è un insieme di  $n \times p$  elementi disposti in  $n$  righe e  $p$  colonne. L'elemento situato all'incrocio tra la riga  $i$  e la colonna  $j$  viene indicato con  $a_{ij}$ . Ciascuna colonna (contenente  $n$  elementi) può essere vista come un punto nello spazio  $R^n$ . Viceversa, ciascuna delle  $n$  righe (contenente  $p$  elementi) può essere vista come un punto nello spazio  $R^p$ . Una matrice  $n \times p$  è anche un insieme di  $p$  vettori colonna (oppure di  $n$  vettori riga). Si chiama scalare un vettore di un solo elemento.

Una matrice  $A$  di dimensione  $n \times p$  si dice quadrata se  $n = p$ . Data una matrice  $A$  di dimensione  $n \times p$  la sua trasposta (di dimensione  $p \times n$ ) si ottiene da  $A$  scambiando tra loro le righe con le colonne. Se  $A' = A$  allora la matrice si dice simmetrica.

Se  $c$  è uno scalare uguale e  $A$  e  $B$  sono due matrici le operazioni più comuni sono:

1.  $c * A$ : moltiplicazione di uno scalare per una matrice;
2.  $A'$ : trasposta di  $A$ ;
3.  $A + B$ : somma elemento per elemento;