7.4.1 Euclidean distance

Euclidean distance is mainly used to calculate distances. The distance between two points in the plane with coordinates (x, y) and (a, b) according to the Euclidean distance formula is given by:

$$\text{Euclidean dist}((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2}$$

For example, the (Euclidean) distance between points (-2, 2) and (2, -1) is calculated as

$$\begin{aligned}\text{Euclidean dist}((-2, 2), (2, -1)) &= \sqrt{(-2 - (2))^2 + (2 - (-1))^2} \\ &= \sqrt{(-4)^2 + (3)^2} \\ &= \sqrt{16 + 9} \\ &= \sqrt{25} \\ &= 5\end{aligned}$$

Table 7.1 Data to calculate Euclidean distances among three persons

	Variable 1	*Variable 2*
Person 1	30	70
Person 2	40	54
Person 3	80	50

Using the formula of Euclidean distance, we can calculate the similarity distance among persons.

The calculation for the distance between person 1 and 2 is:

$$\begin{aligned}\text{Euclidean dist}((30, 70), (40, 54)) &= \sqrt{(30 - 40)^2 + (70 - 54)^2} \\ &= \sqrt{(-10)^2 + (16)^2} \\ &= \sqrt{100+256} \\ &= \sqrt{356} \\ &= 18.86\end{aligned}$$

The calculation for the distance between person 1 and 3 is:

$$\begin{aligned}\text{Euclidean dist}((30, 70), (80, 50)) &= \sqrt{(30 - 80)^2 + (70 - 50)^2} \\ &= \sqrt{(-50)^2 + (20)^2} \\ &= \sqrt{2500 + 400} \\ &= \sqrt{2900} \\ &= 53.85\end{aligned}$$

The calculation for the distance between person 2 and 3 is:

$$\begin{aligned}\text{Euclidean dist}((40, 54), (80, 50)) &= \sqrt{(40 - 80)^2 + (54 - 50)^2} \\ &= \sqrt{(-40)^2 + (4)^2} \\ &= \sqrt{1600+16} \\ &= \sqrt{1616} \\ &= 40.19\end{aligned}$$

7.4.2 Manhattan distance

Manhattan distance is also called L1-distance. It is defined as the sum of the lengths of the projections of the line segment between the two points on the coordinate axes.

For example, the distance between two points in the plane with coordinates (x, y) and (a, b) according to the Manhattan distance formula, is given by:

$$\text{Manhattan dist}((x, y), (a, b)) = |x - a| + |y - b|$$

Let's do the calculations for finding the Manhattan distance among the same three persons, on the basis of their scores on two variables as shown in Table 7.1.

Using the formula of Manhattan distance, we can calculate the similarity distance among persons.

The calculation for the distance between person 1 and 2 is:

$$\begin{aligned}\text{Manhattan dist}((30, 70), (40, 54)) &= |30 - 40| + |70 - 54| \\ &= |-10| + |16| \\ &= 10+16 \\ &= 26\end{aligned}$$

The calculation for the distance between person 1 and 3 is:

$$\begin{aligned}\text{Manhattan dist}((30, 70), (80, 50)) &= |30 - 80| + |70 - 50| \\ &= 50+20 \\ &= 70\end{aligned}$$

The calculation for the distance between person 2 and 3 is:

$$\begin{aligned}\text{Manhattan dist}((40, 54), (80, 50)) &= |40 - 80| + |54 - 50| \\ &= 40+4 = 44\end{aligned}$$

This indicates that the persons 1 and 2 are most similar while person 1 and person 3 are most dissimilar and it produces the same conclusion as Euclidean distance.

Manhattan distance is also called city block distance because like Manhattan, it is the distance a car would drive in a city laid out in square blocks. In Manhattan city, one-way, oblique streets and real streets only exist at the edges of blocks. The Manhattan distance can be represented as shown in Figure 7.3.