Chapter 3

Norm and distance

In this chapter we focus on the norm of a vector, a measure of its magnitude, and on related concepts like distance, angle, standard deviation, and correlation.

3.1 Norm

The *Euclidean norm* of an n-vector x (named after the Greek mathematician Euclid), denoted $\|x\|$, is the squareroot of the sum of the squares of its elements,

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

The Euclidean norm can also be expressed as the squareroot of the inner product of the vector with itself, *i.e.*, $\|x\| = \sqrt{x^T x}$.

The Euclidean norm is sometimes written with a subscript 2, as $\|x\|_2$. (The subscript 2 indicates that the entries of x are raised to the second power.) Other less widely used terms for the Euclidean norm of a vector are the *magnitude*, or *length*, of a vector. (The term *length* should be avoided, since it is also often used to refer to the dimension of the vector.) We use the same notation for the norms of vectors of different dimensions.

As simple examples, we have

$$\left\| \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \right\| = \sqrt{9} = 3, \qquad \left\| \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\| = 1.$$

When x is a scalar, *i.e.*, a 1-vector, the Euclidean norm is the same as the absolute value of x. Indeed, the Euclidean norm can be considered a generalization or extension of the absolute value or magnitude, that applies to vectors. The double bar notation is meant to suggest this. Like the absolute value of a number, the norm of a vector is a (numerical) measure of its magnitude. We say a vector is *small* if its norm is a small number, and we say it is *large* if its norm is a large number. (The numerical values of the norm that qualify for small or large depend on the particular application and context.)

Properties of norm. Some important properties of the Euclidean norm are given below. Here x and y are vectors of the same size, and β is a scalar.

- *Nonnegative homogeneity.* $\|\beta x\| = |\beta|\|x\|$. Multiplying a vector by a scalar multiplies the norm by the absolute value of the scalar.
- *Triangle inequality.* $\|x+y\| \leq \|x\|+\|y\|$. The Euclidean norm of a sum of two vectors is no more than the sum of their norms. (The name of this property will be explained later.) Another name for this inequality is *subadditivity*.
- *Nonnegativity.* $\|x\| \geq 0$.
- *Definiteness.* $\|x\| = 0$ only if $x = 0$.

The last two properties together, which state that the norm is always nonnegative, and zero only when the vector is zero, are called *positive definiteness*. The first, third, and fourth properties are easy to show directly from the definition of the norm. As an example, let's verify the definiteness property. If $\|x\| = 0$, then we also have $\|x\|^2 = 0$, which means that $x_1^2 + \cdots + x_n^2 = 0$. This is a sum of n nonnegative numbers, which is zero. We can conclude that each of the n numbers is zero, since if any of them were nonzero the sum would be positive. So we conclude that $x_i^2 = 0$ for $i = 1, \ldots, n$, and therefore $x_i = 0$ for $i = 1, \ldots, n$; and thus, $x = 0$. Establishing the second property, the triangle inequality, is not as easy; we will give a derivation on page 57.

General norms. Any real-valued function of an n-vector that satisfies the four properties listed above is called a (general) norm. But in this book we will only use the Euclidean norm, so from now on, we refer to the Euclidean norm as the norm. (See exercise 3.5, which describes some other useful norms.)

Root-mean-square value. The norm is related to the *root-mean-square* (RMS) value of an n-vector x, defined as

$$\mathbf{rms}(x) = \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}} = \frac{\|x\|}{\sqrt{n}}.$$

The argument of the squareroot in the middle expression is called the *mean square* value of x, denoted $\mathbf{ms}(x)$, and the RMS value is the squareroot of the mean square value. The RMS value of a vector x is useful when comparing norms of vectors with different dimensions; the RMS value tells us what a 'typical' value of $|x_i|$ is. For example, the norm of $\mathbf{1}$, the n-vector of all ones, is $\sqrt{n}$, but its RMS value is 1, independent of n. More generally, if all the entries of a vector are the same, say, α, then the RMS value of the vector is $|\alpha|$.

Norm of a sum. A useful formula for the norm of the sum of two vectors x and y is

$$\|x + y\| = \sqrt{\|x\|^2 + 2x^T y + \|y\|^2}. \tag{3.1}$$

To derive this formula, we start with the square of the norm of $x+y$ and use various properties of the inner product:

$$\begin{aligned}\|x+y\|^2 &= (x+y)^T(x+y)\\ &= x^Tx + x^Ty + y^Tx + y^Ty\\ &= \|x\|^2 + 2x^Ty + \|y\|^2.\end{aligned}$$

Taking the squareroot of both sides yields the formula (3.1) above. In the first line, we use the definition of the norm. In the second line, we expand the inner product. In the fourth line we use the definition of the norm, and the fact that $x^Ty = y^Tx$. Some other identities relating norms, sums, and inner products of vectors are explored in exercise 3.4.

Norm of block vectors. The norm-squared of a stacked vector is the sum of the norm-squared values of its subvectors. For example, with $d = (a, b, c)$ (where a, b, and c are vectors), we have

$$\|d\|^2 = d^Td = a^Ta + b^Tb + c^Tc = \|a\|^2 + \|b\|^2 + \|c\|^2.$$

This idea is often used in reverse, to express the sum of the norm-squared values of some vectors as the norm-square value of a block vector formed from them.

We can write the equality above in terms of norms as

$$\|(a,b,c)\| = \sqrt{\|a\|^2 + \|b\|^2 + \|c\|^2} = \|(\|a\|, \|b\|, \|c\|)\|.$$

In words: The norm of a stacked vector is the norm of the vector formed from the norms of the subvectors. The right-hand side of the equation above should be carefully read. The outer norm symbols enclose a 3-vector, with (scalar) entries $\|a\|$, $\|b\|$, and $\|c\|$.

Chebyshev inequality. Suppose that x is an n-vector, and that k of its entries satisfy $|x_i| \geq a$, where $a > 0$. Then k of its entries satisfy $x_i^2 \geq a^2$. It follows that

$$\|x\|^2 = x_1^2 + \cdots + x_n^2 \geq ka^2,$$

since k of the numbers in the sum are at least a^2, and the other $n-k$ numbers are nonnegative. We can conclude that $k \leq \|x\|^2/a^2$, which is called the *Chebyshev inequality*, after the mathematician Pafnuty Chebyshev. When $\|x\|^2/a^2 \geq n$, the inequality tells us nothing, since we always have $k \leq n$. In other cases it limits the number of entries in a vector that can be large. For $a > \|x\|$, the inequality is $k \leq \|x\|^2/a^2 < 1$, so we conclude that $k = 0$ (since k is an integer). In other words, no entry of a vector can be larger in magnitude than the norm of the vector.

The Chebyshev inequality is easier to interpret in terms of the RMS value of a vector. We can write it as

$$\frac{k}{n} \leq \left(\frac{\mathbf{rms}(x)}{a}\right)^2, \tag{3.2}$$

where k is, as above, the number of entries of x with absolute value at least a. The left-hand side is the fraction of entries of the vector that are at least a in absolute

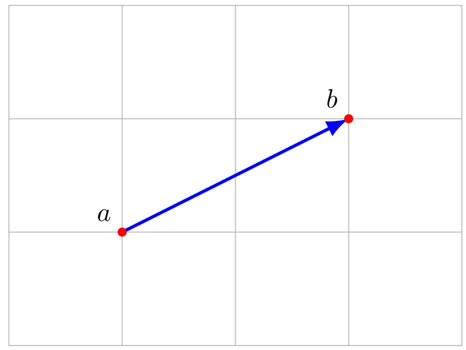

Figure 3.1 The norm of the displacement $b - a$ is the distance between the points with coordinates a and b.

value. The right-hand side is the inverse square of the ratio of a to $\mathbf{rms}(x)$. It says, for example, that no more than $1/25 = 4\%$ of the entries of a vector can exceed its RMS value by more than a factor of 5. The Chebyshev inequality partially justifies the idea that the RMS value of a vector gives an idea of the size of a typical entry: It states that not too many of the entries of a vector can be much bigger (in absolute value) than its RMS value. (A converse statement can also be made: At least one entry of a vector has absolute value as large as the RMS value of the vector; see exercise 3.8.)

3.2 Distance

Euclidean distance. We can use the norm to define the *Euclidean distance* between two vectors a and b as the norm of their difference:

$$\mathbf{dist}(a, b) = \|a - b\|.$$

For one, two, and three dimensions, this distance is exactly the usual distance between points with coordinates a and b, as illustrated in figure 3.1. But the Euclidean distance is defined for vectors of any dimension; we can refer to the distance between two vectors of dimension 100. Since we only use the Euclidean norm in this book, we will refer to the Euclidean distance between vectors as, simply, the distance between the vectors. If a and b are n-vectors, we refer to the RMS value of the difference, $\|a - b\|/\sqrt{n}$, as the *RMS deviation* between the two vectors.

When the distance between two n-vectors x and y is small, we say they are 'close' or 'nearby', and when the distance $\|x - y\|$ is large, we say they are 'far'. The particular numerical values of $\|x - y\|$ that correspond to 'close' or 'far' depend on the particular application.