

arithm. As a result, the maximum likelihood solution for the parameters no longer has a closed-form analytical solution. One approach for maximizing the likelihood function is to use iterative numerical optimization techniques. Alternatively, we can employ a powerful framework called *expectation maximization*, which has wide applicability to a variety of different deep generative models.

3.3. Periodic Variables

Although Gaussian distributions are of great practical significance, both in their own right and as building blocks for more complex probabilistic models, there are situations in which they are inappropriate as density models for continuous variables. One important case, which arises in practical applications, is that of periodic variables.

An example of a periodic variable is the wind direction at a particular geographical location. We might, for instance, measure the wind direction at multiple locations and wish to summarize this data using a parametric distribution. Another example is calendar time, where we may be interested in modelling quantities that are believed to be periodic over 24 hours or over an annual cycle. Such quantities can conveniently be represented using an angular (polar) coordinate $0 \leq \theta < 2\pi$.

We might be tempted to treat periodic variables by choosing some direction as the origin and then applying a conventional distribution such as the Gaussian. Such an approach, however, would give results that were strongly dependent on the arbitrary choice of origin. Suppose, for instance, that we have two observations at $\theta_1 = 1^\circ$ and $\theta_2 = 359^\circ$, and we model them using a standard univariate Gaussian distribution. If we place the origin at 0° , then the sample mean of this data set will be 180° with standard deviation 179° , whereas if we place the origin at 180° , then the mean will be 0° and the standard deviation will be 1° . We clearly need to develop a special approach for periodic variables.

3.3.1 Von Mises distribution

Let us consider the problem of evaluating the mean of a set of observations $\mathcal{D} = \{\theta_1, \dots, \theta_N\}$ of a periodic variable θ where θ is measured in radians. We have already seen that the simple average $(\theta_1 + \dots + \theta_N)/N$ will be strongly coordinate dependent. To find an invariant measure of the mean, note that the observations can be viewed as points on the unit circle and can therefore be described instead by two-dimensional unit vectors $\mathbf{x}_1, \dots, \mathbf{x}_N$ where $\|\mathbf{x}_n\| = 1$ for $n = 1, \dots, N$, as illustrated in Figure 3.9. We can average the vectors $\{\mathbf{x}_n\}$ instead to give

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (3.117)$$

and then find the corresponding angle $\bar{\theta}$ of this average. Clearly, this definition will ensure that the location of the mean is independent of the origin of the angular coordinate. Note that $\bar{\mathbf{x}}$ will typically lie inside the unit circle. The Cartesian coordinates