1

Key Points

This chapter introduces mathematical concepts needed in the remainder. Readers with background in statistics or machine learning may find that they can skim through it. In fact, we recommend that readers skip this chapter on their first, high-level pass through this text, as the first conceptual arguments for Probabilistic Numerics will arrive in Chapter II. However, the mathematical arguments made in later chapters require the following key concepts, which must be developed here first:

§ 2 Probabilities provide the formal framework for reasoning and inference in the presence of uncertainty.

§ 3 The notion of computation as probabilistic inference is not restricted to one kind of probability distribution; but Gaussian distributions play a central role in inference on continuous-valued variables due to their convenient algebraic properties.

§ 4 *Regression* – inference on a function from observations of its values at a finite number of points is an internal operation of all numerical methods and arguably a low-level numerical algorithm in itself. But regression is also a central task in machine learning. We develop the canonical mathematical tools for probabilistic regression:

§ 4.1 Gaussian uncertainty over the weights of a set of basis functions allows inference on both linear and nonlinear functions within a finite-dimensional space of hypotheses. The choice of basis functions for this task is essentially unlimited and has wide-ranging effects on the inference process.

https://doi.org/10.1017/9781316681411.005 Published online by Cambridge University Press

§ 4.2 *Gaussian processes* extend the notion of a Gaussian distribution to infinite-dimensional objects, such as functions.

§ 4.3 In the Gaussian case, the probabilistic view is very closely related to other frameworks of inference and approximation, in particular to *least-squares estimation*. This connection will be used in later chapters of this texts to connect classic and probabilistic numerical methods. In fact, readers with a background in interpolation/scattered data approximation may interpret this section as covering a first example of building a probabilistic interpretation for these numerical tasks. In contrast, from the perspective of machine learning and statistics, regression is not a computational task, but a principal form of learning. This difference in viewpoints thus highlights the fundamental similarity of computation and learning once again.

§ 4.4 Gaussian process models allow inference on *derivatives* of functions from observations of the function, and vice versa. This will be relevant in all domains of numerical computation, in particular in integration, optimisation, and the solution of differential equations.

§ 5 *Gauss–Markov processes* are a class of Gaussian process models on univariate domains whose "finite memory" allows inference at cost linear in the number of observations. The inference process is packaged into algorithms known as *filters* and *smoothers*. The dynamics of the associated stochastic process are captured by *linear stochastic differential equations*.

§ 6 Conjugate priors for the *hyperparameters* of Gaussian models allow inference on the prior mean and covariance at low computational cost.