5.2.3 向量与矩阵的运算及期望值

上一节提到了我们将借助图形来理解协方差，不少读者也许已经迫不及待想要一探究竟了。不过在此之前，我们还需要学习一些准备知识。这些内容虽然基础，但对于准确理解之后的讲解不可或缺。我们争取在本节把所有这些内容全部解决。

为了处理取值随机的向量与取值随机的矩阵，问答专栏5.1引入了值为向量与矩阵的随机变量。之后，我们自然会在算式中遇到向量与矩阵的运算以及相关的期望值。本书并非故意讨论这些复杂的计算，向量与矩阵的运算在统计分析、模式识别和信号处理等领域都有着广泛的应用。因此，本节将帮助读者练习这方面的计算。

设 $\boldsymbol{X}$ 为 n 元随机列向量（如果觉得 n 元不容易理解，也可以使用二元或三元等具体的值）。前面已经介绍过，$\boldsymbol{X}$ 的期望值由其中每一个元素的期望值决定。

$$\mathrm{E}[\boldsymbol{X}] = \mathrm{E}\left[\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}\right] \equiv \begin{pmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \\ \vdots \\ \mathrm{E}[X_n] \end{pmatrix}$$

根据该定义不难看出，常量 c、常向量 $\boldsymbol{a}$ 或向量值随机变量 $\boldsymbol{Y}$ 与该期望值存在以下关系，与数值的期望值没有区别（设 $\boldsymbol{a}$ 与 $\boldsymbol{Y}$ 的维数与 $\boldsymbol{X}$ 相同）。

$$\mathrm{E}[c\boldsymbol{X}] = c\mathrm{E}[\boldsymbol{X}], \quad \mathrm{E}[\boldsymbol{X}+\boldsymbol{a}] = \mathrm{E}[\boldsymbol{X}]+\boldsymbol{a}, \quad \mathrm{E}[\boldsymbol{X}+\boldsymbol{Y}] = \mathrm{E}[\boldsymbol{X}]+\mathrm{E}[\boldsymbol{Y}], \quad \mathrm{E}[\boldsymbol{a}] = \boldsymbol{a}$$

假设有一个取值恒定的列向量 $\boldsymbol{a}$，$\boldsymbol{a}$ 与 $\boldsymbol{X}$ 的内积的期望值如下。

$$\mathrm{E}[\boldsymbol{a}\cdot\boldsymbol{X}] = \mathrm{E}[\boldsymbol{a}^T\boldsymbol{X}] = \boldsymbol{a}^T\mathrm{E}[\boldsymbol{X}] = \boldsymbol{a}\cdot\mathrm{E}[\boldsymbol{X}]$$

如果读者无法理解这部分计算，请参见附录A.6。当然，上式成立的前提是 $\boldsymbol{a}$ 与 $\boldsymbol{X}$ 维数相同（否则不满足内积的条件）。此外，请读者注意，行向量与列向量的乘积是一个数字[①]。只要写出其中的元素，就能很容易证明上式成立。事实上，对于 $\boldsymbol{X} \equiv (X_1, \cdots, X_n)^T$ 与 $\boldsymbol{a} \equiv (a_1, \cdots, a_n)^T$，我们可以对式子做如下变形，证明两者的值相同。

① 我们在此默认读者理解这一规定，如果有任何问题，请复习线性代数的相关章节。我们建议初学者在学习这部分内容时，仔细判断式中的每一个元素，确认它们究竟是数字还是向量抑或是矩阵。如果不能立即指出式中的元素表示的含义，请不要急着继续阅读，而应静下心来认真领悟。如果没能深入理解这些元素的含义，就很容易发生错误。这也是本章的一个难点。