frames of speech. Considering that each frame spans about $10ms$, then, this would be equivalent to having $27,673.5s$ or 461.2 minutes!

Of course it is quite difficult to come up with 461 minutes of speech for training. There are different remedies to this problem. One is to do away with the diagonal elements of the covariance matrix and only consider the variances. This will reduce the number of required parameters to

$$\begin{aligned}\dim(\boldsymbol{\varphi}) &= \Gamma(2D+1)-1\\ &= 256\,(2\times 45+1)-1\\ &= 23,295 \qquad (13.80)\end{aligned}$$

which is 91.6% less than the number of parameters needed in a full covariance model. The number of minutes of audio for estimating these parameters would then only be 38.83 minutes which is much more practical.

However, we usually do not use the data from only one speaker to estimate the parameters of the whole set of parameters in a model. Generally, a large training set is used, which consists of hundreds and maybe even thousands of speakers to compute the basic model parameters. This is sometimes called the *speaker independent model* and in some circumstances it is referred to as the *universal model*, *universal background model*, or simply *background model*. As we will see in chapter 21, the individual target speaker's enrollment audio is generally used to adapt new model parameters from this *speaker independent model*.

Another approach is to reduce the dimensionality of the data. This would mean the reduction of D which greatly affects the total number of free parameters. In Chapter 12 we discussed a few techniques designed to optimize the amount of information which is preserved by reducing the dimensionality of the free parameters in the system. Some such techniques are *PCA*, *LDA*, and *FA*. We shall see more about such dimensionality reduction as applied specifically to the speaker recognition problems, such as *joint factor analysis* (Section 16.4.2) and *nuisance attribute projection* (Section 15.8).

In Chapter 16, we will speak more about the *speaker model* and how it is devised. In the chapters following Chapter 16, more detail will be given regarding different aspects of model and parameter optimization.

13.8 Practical Issues

In this section, we will examine a few practical issues involved in the successful training of *HMM* and *GMM*. The first and one of the foremost problems with training a statistical system, regardless of whether it is an HMM or a GMM is the short-