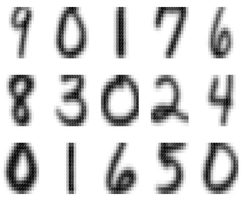

Figure 10.4 Left to Right, Top to Bottom. Patterns corresponding to the first basis functions. Note that 9 out of 10 digits are chosen among the first 10 patterns and that all patterns are sufficiently different (or shifted).

performance.

A practical use of SGMA lies in the fact that it allows us to find increasingly accurate, yet sparse approximations $\tilde{K}$ which can subsequently be used in optimization algorithms as replacements for K. It is also worth noting that the approximation and training algorithms can be coupled to obtain fast methods that do not require computation of K. Many methods find a dense expansion first and only subsequently employ a reduced set algorithm (Chapter 18) to find a more compact representation.

10.3 Interior Point Algorithms

Interior point algorithms are some of the most reliable and accurate optimization techniques and are the method of choice for small and moderately sized problems. We will discuss approximations applicable to large-scale problems in Sections 10.4, 10.5 and 10.6. We assume that the reader is familiar with the basic notions of interior point algorithms. Details can be found in Section 6.4 and references therein. In this section we focus on Support Vector specific details.

In order to deal with optimization problems which have both equality constraints and box constrained variables, we need to extend the notation of (6.72) slightly. The following optimization problem is general enough to cover classification, regression, and novelty detection:

$$\begin{aligned} \underset{\alpha,t}{\text{minimize}} \quad & \tfrac{1}{2}\alpha^\top Q\alpha + c^\top \alpha \\ \text{subject to} \quad & A\alpha = d \\ & \{0 \le \alpha \le u\} \text{ or } \{\alpha + t = u \text{ and } \alpha, t \ge 0\}. \end{aligned} \tag{10.43}$$

Here Q is a square matrix, typically of size $m \times m$ or $(2m) \times (2m)$, c, α, t, u are vectors of the same dimension, and A is a corresponding rectangular matrix. The dual can be found to be

$$\begin{aligned} \underset{\alpha,s,z,h}{\text{maximize}} \quad & -\tfrac{1}{2}\alpha^\top Q\alpha + d^\top h - u^\top s \\ \text{subject to} \quad & Q\alpha + c - A^\top h + s - z = 0 \\ & s, z \ge 0 \text{ and } h \text{ free} \end{aligned} \tag{10.44}$$

Furthermore, we have the Karush-Kuhn-Tucker (KKT) conditions

$$\alpha_i z_i = 0 \text{ and } s_i t_i = 0 \text{ for all } i \in [m] \tag{10.45}$$