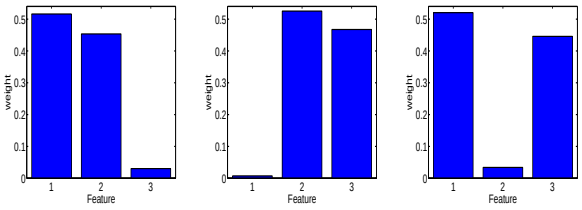
Appendices

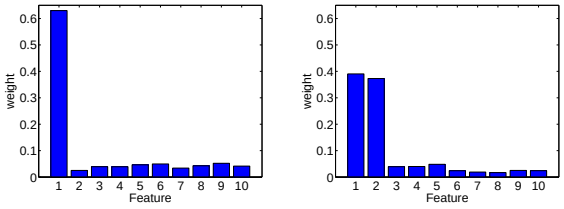
Appendix A Results of Simulated Data

A.1 Analysis of subspace exploration. WAMS succeeds in discovering the subspace of each cluster. The weight vectors of the clusters obtained by WAMS ($k = 50$) on the three toy examples are shown in Figs. A.1, A.2, and A.3, respectively. In Fig. A.1, cluster 1 (corresponding to class 1 in Toy1; '#150' indicates that cluster 1 contains 150 points) has larger weights on the 1st and 2nd features, and a very small weight on the 3rd feature. This means that cluster 1 exists in the subspace spanned by the first two features. This estimation perfectly matches the underlying distribution that has generated the data (as described in Fig. 1). Moreover, we can see from Fig. 1 (b) that the dispersion of the 1st feature (x-axis) is lower than the one of the 2nd feature (y-axis). This is the reason why the 1st feature receives a weight larger than the weight of the 2nd feature, as seen in Fig. A.1 (a). Similar results were obtained for clusters 2 and 3, as shown in Fig. A.1. In Figs. A.2 and A.3, for cluster 1 (corresponding to class 1 in Toy2 and Toy3) the largest weight is assigned to the 1st feature. Looking at Fig. 2, the dispersion of the data along the y-axis is much larger than on the x-axis; thus the y-axis can be considered as an irrelevant feature for class 1. For this reason it receives a very small weight, along with the other noisy features. Comparable large weights are assigned to the first two features in cluster 2 on both Toy2 and Toy3 (as expected from the distribution of class 2 in Fig. 2).



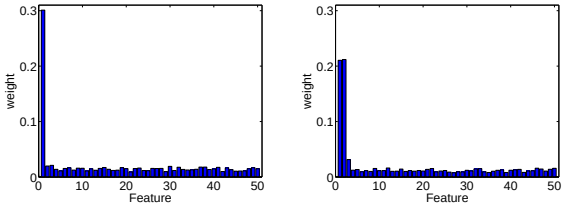
(a) cluster 1(#150) (b) cluster 2(#150) (c) cluster 3(#150)

Figure A.1: Weight distribution of WAMS on Toy1 ($k = 50$)



(a) cluster 1(#150) (b) cluster 2(#150)

Figure A.2: Weight distribution of WAMS on Toy2 ($k = 50$)



(a) cluster 1(#145) (b) cluster 2(#155)

Figure A.3: Weight distribution of WAMS on Toy3 ($k = 50$)