28 MARCH 2007
HANDOUT

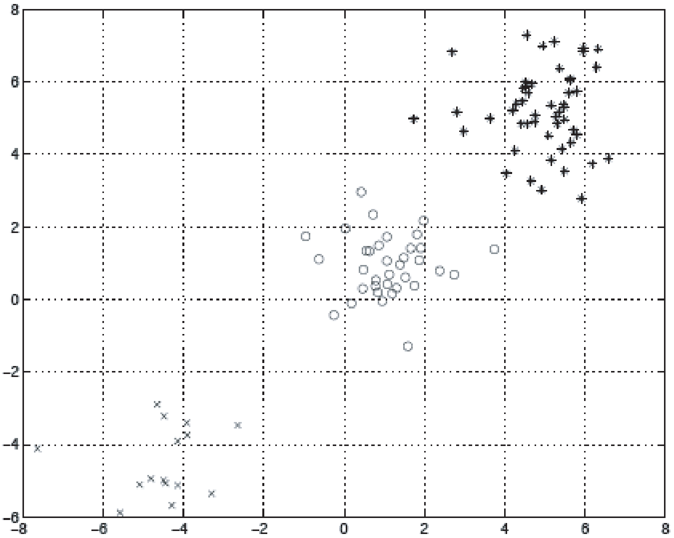

FIG. 0.1. *There are 100 data points. The 50 '+' points are normally distributed* $(0, 1)$ *perturbations added to the point* $(5, 5)$. *The next 35 '○' are normally distributed* $(0, 1)$ *perturbations added to the point* $(1, 1)$. *The last 15 '×' are normally distributed* $(0, 1)$ *perturbations added to the point* $(-1, -1)$.

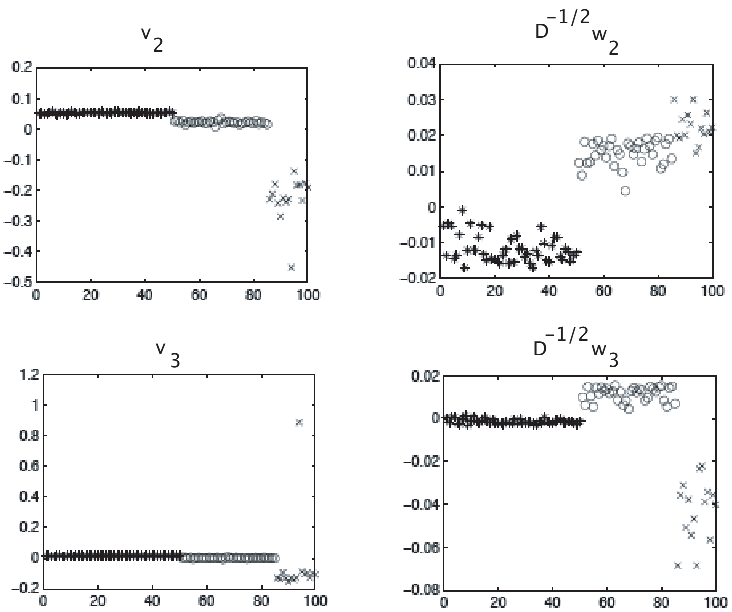

FIG. 0.2. *The unnormalized Fiedler vector* v_2 *identifies the three natural clusters. Also, it notices the outlier. Since the partition is recognized completely,* v_3 *doesn't do much except recognize the outlier. The normalized Fiedler vector* $D^{-1/2}w_2$ *separates the data into 2 sets, each with 50 points (by clustering the '○' and '×' points together). The third normalized eigenvector* $D^{-1/2}w_3$ *separates the second cluster into '○' and '×' separately.*

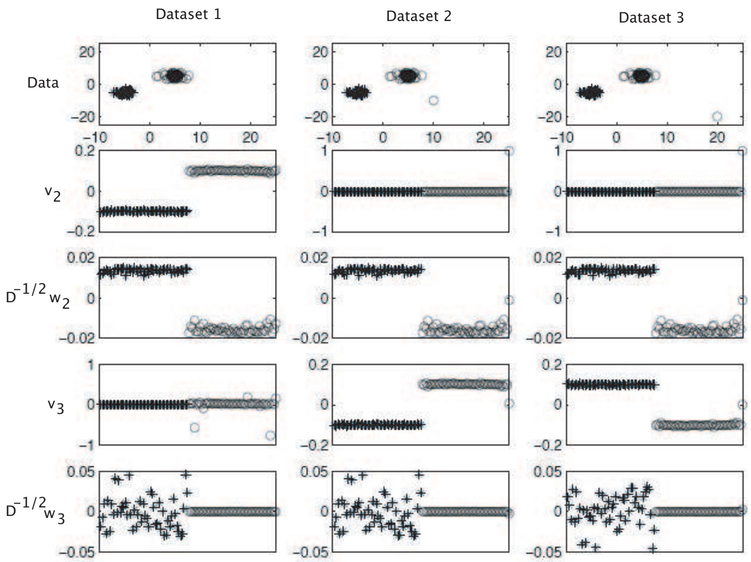

FIG. 0.3. ***Dataset 1***: *There are 2 clusters of data, 50 '+' points and 50 '∘' points. The* 2^{nd} *eigenvector (both normalized and unnormalized) partitions the data into the 2 natural subsets. Thus the* 3^{rd} *eigenvector is irrelevant.* ***Dataset 2***: *This is the same as Dataset 1, except with one outlier. The unnormalized eigenvector* v_2 *separates the outlier and the remainder of the data. Then* v_3 *separates the non-outlier partition into the 2 original partitions. For the normalized eigenvector* $D^{-1/2}w_2$, *we get the 2 original partitions and a third partition just containing the outlier. Thus,* $D^{-1/2}w_3$ *tells us nothing.* ***Dataset 3***: *This is the same as Dataset 2, but the outlier is farther from the other data. As with dataset 2, the unnormalized eigenvector* v_2 *separates the outlier and the remainder of the data. Then* v_3 *separates the non-outlier partition into the 2 original partitions. For the normalized eigenvector* $D^{-1/2}w_2$, *we get the 2 original partitions and a third partition just containing the outlier. Thus,* $D^{-1/2}w_3$ *tells us nothing.*

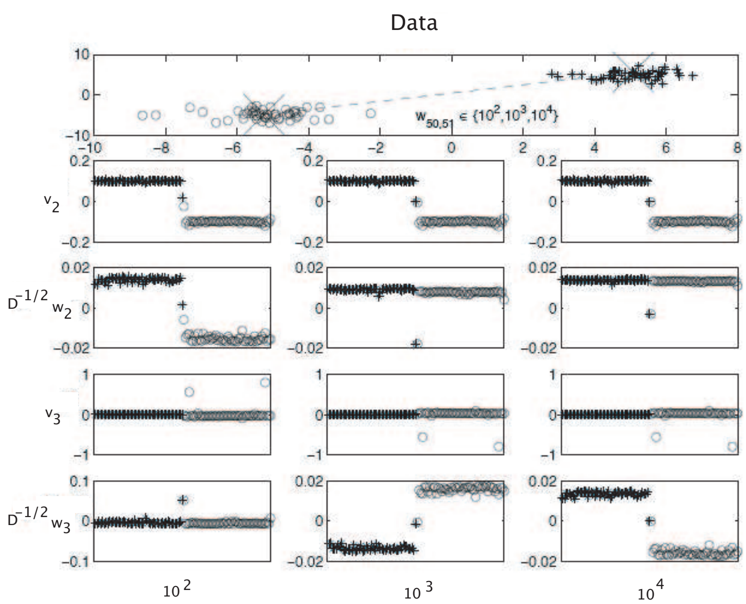

FIG. 0.4. *There are 2 clusters of data, each with 50 data points. To study the effect of a strong weight, the value* $w_{50,51}$ *takes on values* 10^2 *for the first column,* 10^3 *for the second column, and* 10^4 *for the third column. Before this replacement, the largest* w_{ij} *was 12.1. When* $w_{50,51} = 10^2$: *the vectors* v_2 *and* $D^{-1/2}w_2$ *separate the data into the 2 large partitions, and a third partition with* $(50, 51)$. *Thus, in this case* v_3 *and* $D^{-1/2}w_3$ *do not tell us anything. For the other cases,* $w_{50,51} = 10^3$ *and* $w_{50,51} = 10^4$, v_2 *partitions the data into the 2 main sets, and a third with* $(50, 51)$, *and thus* v_3 *tells us nothing. The normalized vector* $D^{-1/2}w_2$ *separates the data into* $(50, 51)$ *and everything else. Then* $D^{-1/2}w_3$ *separates the remaining data into the 2 original partitions.*