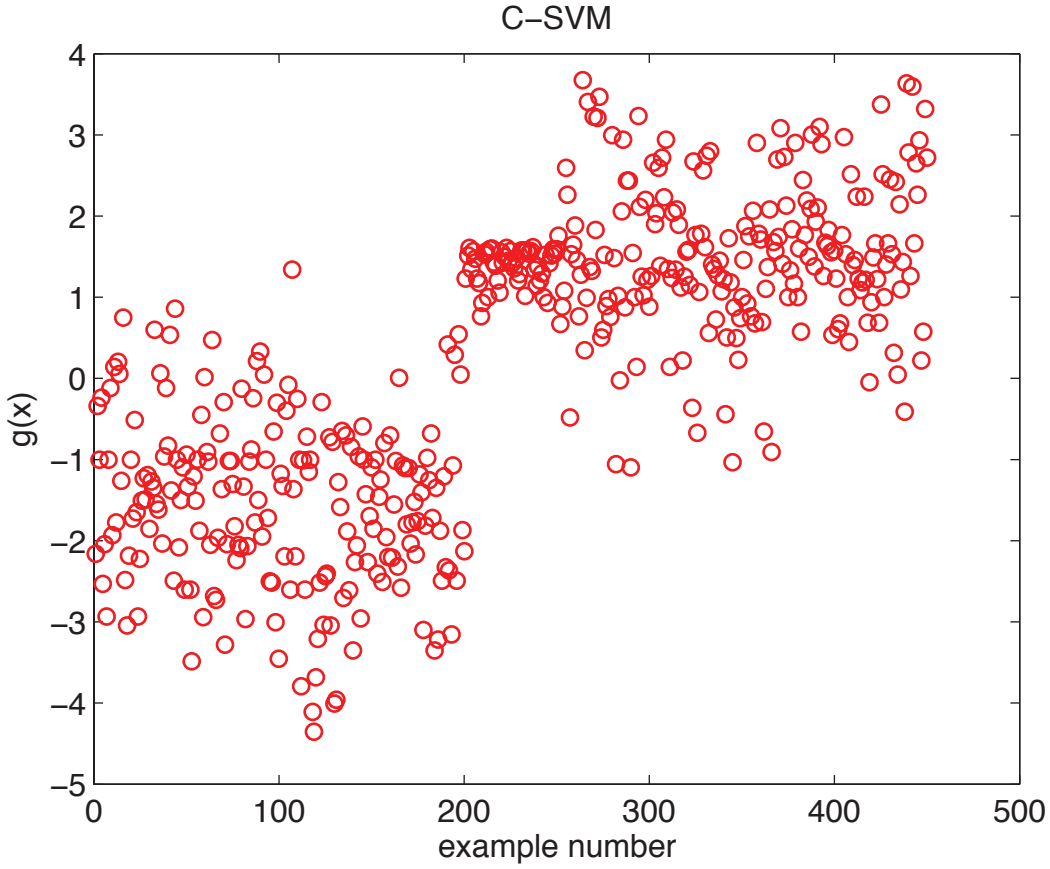
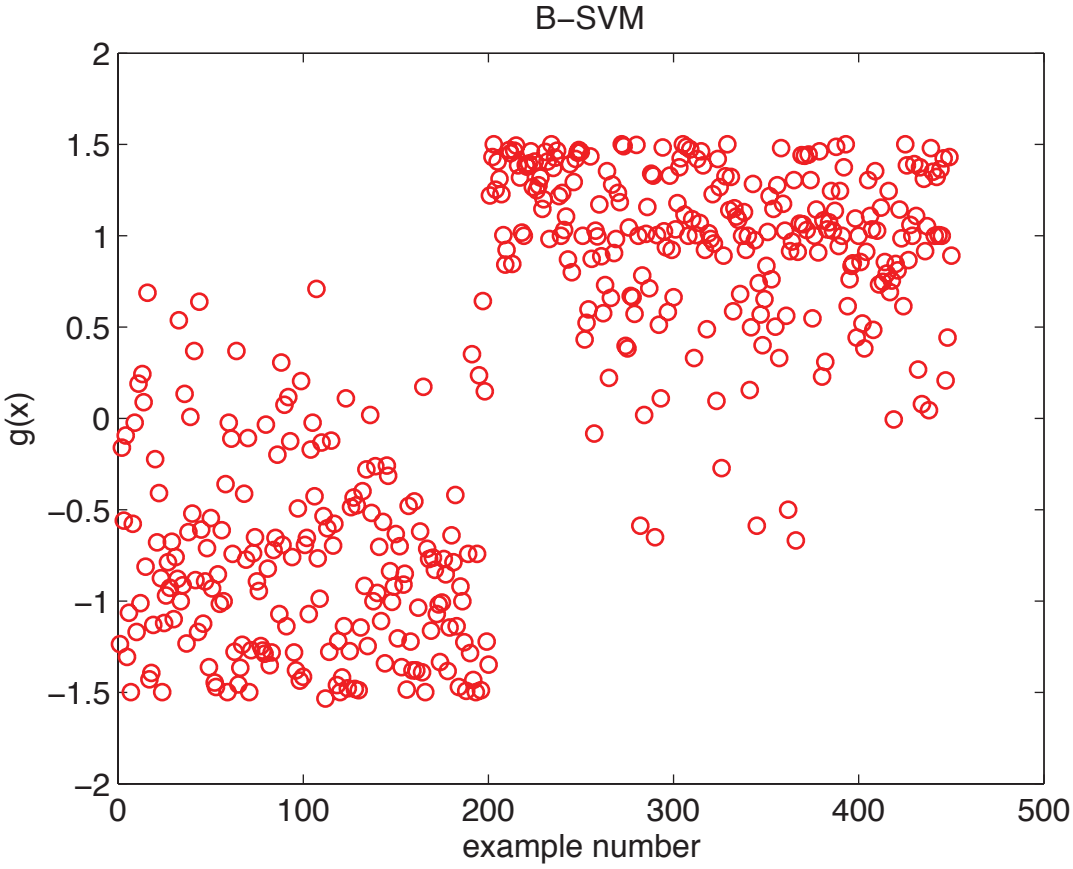
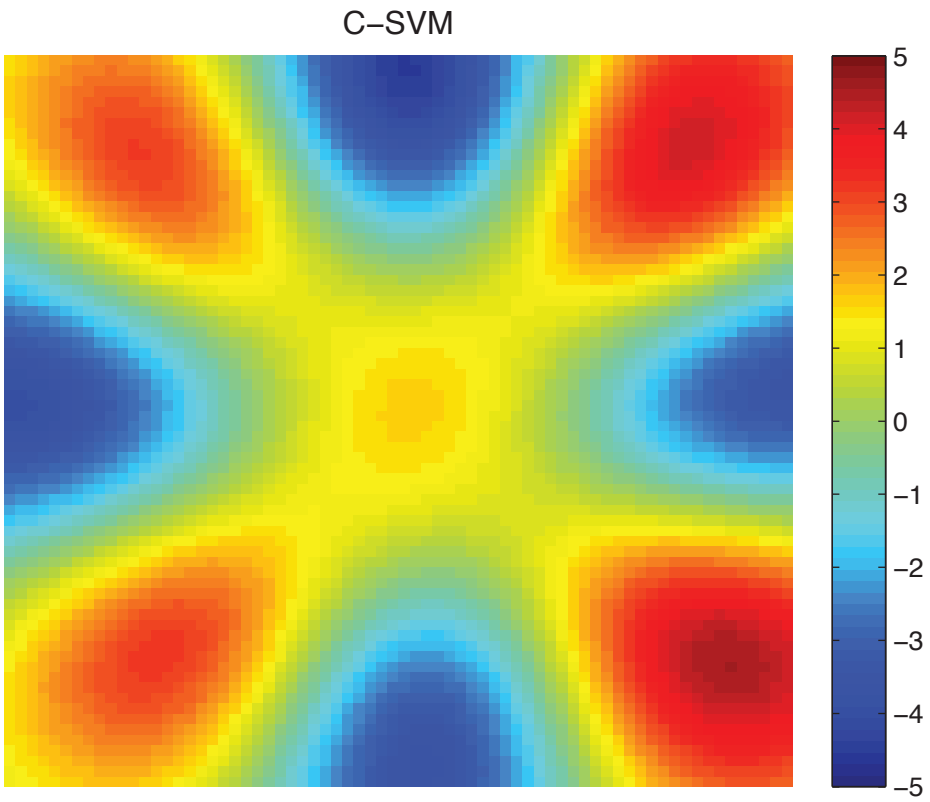


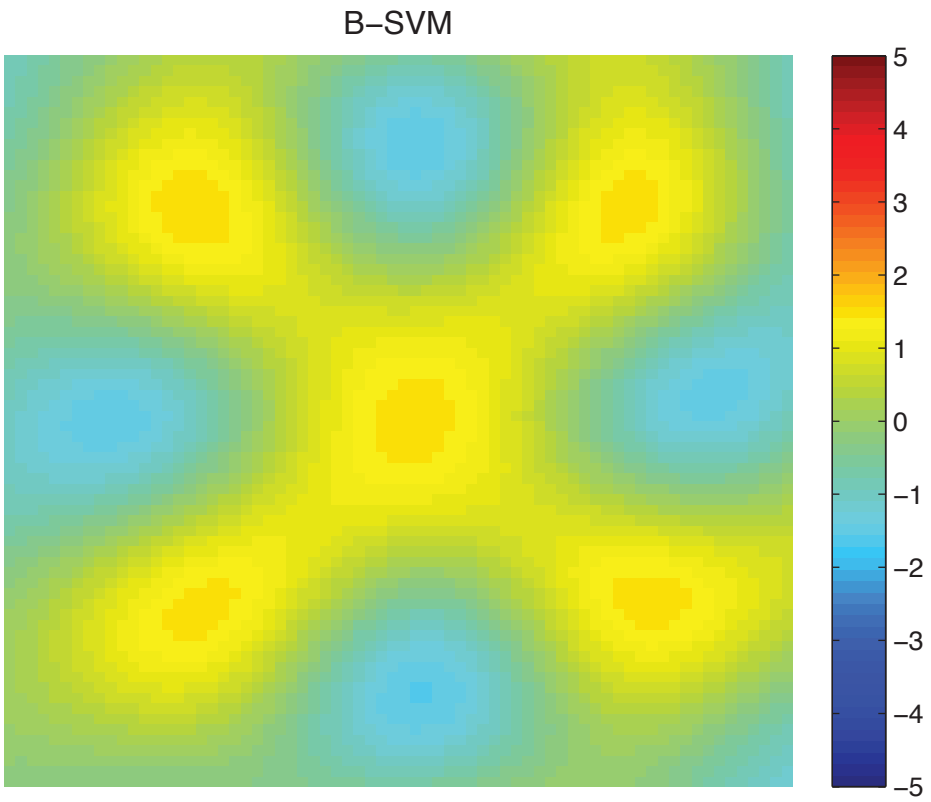
(a) C-SVM $g(\boldsymbol{x})$



(b) B-SVM $g(\boldsymbol{x})$



(c) C-SVM $g(\boldsymbol{x})$ heatmap



(d) B-SVM $g(\boldsymbol{x})$ heatmap

Figure 3: Figure shows decision rule $g(\boldsymbol{x})$ for C-SVM (a) and B-SVM (b). Note that in B-SVM the second penalty term $C_2 \sum_{i=1}^{n} [y_i(\boldsymbol{\beta}^T \boldsymbol{h}(\boldsymbol{x_i}) + \beta_0) - \rho_2]_+$ results in most of the $g(\boldsymbol{x})$ values in the interval $[\rho_1, \rho_2] = [1, 1.5]$. (c) Heat map of the decision rule $g(\boldsymbol{x})$ for C-SVM (d) Heat map of the decision rule $g(\boldsymbol{x})$ for B-SVM. In C-SVM the values of decision rule $g(\boldsymbol{x})$ are unbalanced in Class 1. The central cluster located at $(0, 0)$ in Class 1 gets much smaller $g(\boldsymbol{x})$ values in C-SVM than the rest of the Class 1. In B-SVM, however, all clusters in Class 1 including the one centered at $(0, 0)$ get similar $g(\boldsymbol{x})$ values. This is a result of the second penalty term in the B-SVM objective function.