Semi-supervised Eigenvectors for Locally-biased Learning

Methods

We consider the **Generalized LocalSpectral** problem in Figure 2, i.e., a constrained eigenvalue problem. We seek a vector \( x \in \mathbb{R}^n \) that minimizes \( \frac{1}{2} \langle x^T L_{G} x \rangle \) subject to several constraints; that \( x \) is unit length, that \( x \) is orthogonal to the span of \( Q \) (previous solutions and the all-one vector); and that \( x \) is \( \gamma \)-well-correlated with the input seed set vector \( s \).

**Figure 2:** Left: The usual **GlobalSpectral** partitioning optimization problem, the vector achieving the optimal solution is \( \lambda \gamma \), the leading nontrivial generalized eigenvector of \( L_{G} \) with respect to \( D_{G} \). Middle: The **LocalSpectral** optimization problem, which was originally introduced in [5]. For \( \kappa = 0 \), this coincides with the usual global spectral objective, while for \( \kappa > 0 \), this produces solutions that are biased toward the seed vector \( s \). Right: The **Generalized LocalSpectral** optimization problem we introduce that includes both the locality constraint and a more general orthogonality constraint. Our main algorithm for computing semi-supervised eigenvectors will iteratively compute the solution to **Generalized LocalSpectral** for a sequence of \( Q \) matrices. In all three cases, the optimization variable is \( x \in \mathbb{R}^n \).

The \( k \)th semi-supervised eigenvector is given by \( x_{k} \propto (F^{T} L_{G} = \gamma D_{G} F^{T})^{-1} D_{G} x \) where \( F^{-1} \) is a projection matrix of the form \( F = I - D_{G} Q / Q^{T} D_{G} Q = Q^{T} D_{G}^{-1} Q / Q^{T} D_{G} Q \).

To determine the correct setting of \( \gamma \), it suffices to perform a binary search over the interval \([\max(\lambda_{G}) \gamma, \min(\lambda_{G}) \gamma] \) for the constraint correlation constant is satisfied (\( G \) denotes a connected undirected graph), for more details see [2, 5].

**Results**

We consider the well-studied MNIST dataset containing 60000 training digits and 10000 test digits ranging from 0 to 9. We construct the complete \( 70000 \times 70000 \) k-NN graph with \( k = 10 \) and edge weights given by an exponential kernel. We then evaluate the semi-supervised eigenvectors in a transductive learning setting by disregarding the majority of training labels, and we are interested in discriminating between fours and nines, as these two classes tend to overlap more than other combinations.

Figure 3 visualizes classification results based on the Spectral Graph Transducer (SGT) of [4], see the caption for further details.

Figure 4 demonstrates that classification accuracy is robust with respect to the choice of \( \kappa \) distribution, see the caption for further details.

**References**


