GRAPH LAPLACIAN AND LINEAR SMOOTHER REPORT 2.12

Graph Laplacian & Linear Smoothers
On Nonlinear Dimensionality Reduction, Linear Smoothing and Autoencoding
Matrices obtained from local, spectral methods can be seen as an operator that returns the bias from smoothing a function. Each of the NLDR methods considered construct or approximate a matrix $L = G(I - S)$ where G is a diagonal matrix and S is a linear smoother. In this sense, L measures the bias of the prediction Sf weighted by G. For example, the Diffusion Maps constructs the Nadaraya-Watson kernel smoother $S = D^{-1}K$ where D is the degree matrix and K the kernel matrix. The constructed embedding consists fo the right singular vectors corresponding to the smallest singular values of $L^{DM} = I - S$, which is also generally called the random walk Laplacian.Laplacian Eigenmaps, or the unnormalized graph Laplacian by another name, is $L^{LE} = DL^{DM} = D(I - S)$, is a rescaling of the bias.

- Connection between the spectrum of linear smoothers and spectrum of graph Laplacians?

Smoothers and penalized least squares
Linear Smoothers and Additive Models
There is a class of smoothers that can be characterized as solutions to penalized least-squares, especially in the form:

$$||\boldsymbol{y} - f||^2 + \lambda f^t K f$$

Since the penalization term is in quadratic form, only symmetric matrices K should be considered. Also, the smoothing matrix $S = (I + \lambda K)^{-1}$ given the inverse is well defined. Conversely, given a smoother S, can frame into penalization problem as

$$||\boldsymbol{y} - f||^2 + f^t(S^{-1} - I)f$$

Graph Laplacian & Regularization/Penalization on Manifolds
Manifold Regularization: A geometric Framework for Learning from Labeled and Unlabeled Examples

To incorporate intrinsic geometry of the manifold, can frame a regularization problem as:

$$f^* = argmin_{f \in \mathcal{H}} \frac{1}{l} \sum_{i=1}^{l} V(x_i, y_i, f) + \gamma_A ||f||^2_{\mathcal{H}} + \gamma_I ||f||^2_I$$

where the additional last term reflect the intrinsic structure of the data manifold. A natural choice for $||f||^2_I$ is $\int_{x \in \mathcal{M}} ||\nabla_{\mathcal{M}} f||^2 dP_X(x)$ and by defining a new kernel $\tilde{k}$ on the manifold the above problem can be written as

$$f^* = argmin_{f \in \mathcal{H}} \frac{1}{l} \sum_{i=1}^{l} V(x_i, y_i, f) + \gamma ||f||^2_{\tilde{\mathcal{H}}}$$

Date: April 18, 2021.

and given the Representor Theorem holds and we have $f^* = \sum_{i=1}^n \alpha_i \tilde{k}(x_i, .)$, the penalty can be written in quadratic form as $\gamma \boldsymbol{\alpha}^T \tilde{K} \boldsymbol{\alpha}$ where $\boldsymbol{\alpha}$ is the vector of the coefficients and $\tilde{K}$ is the Gram matrix with the new kernel. For the method in the paper,

Connections between the three approaches:
Consistency results

- Linear smoothers: Nearest-neighbor-type smoothers in Stone (1977), Cubic splines in Cox (1983) and Rice & Rosenblatt (1983), and kernel smoothers in Gasser & Muller (1979).
- Graph Laplacian: Laplacian Eigenmap (Belkin), Diffusion Maps (Coifman & Lafon), Ting, Berry, Trillos ...
- Manifold Regularization: Haven't checked any consistency result yet

Transition between the fields:
When can a method in one approach be constructed/framed from another approach?