Abstracts

Hodge Decomposition and Learning Theory

Steve Smale
(joint work with Nat Smale)

Partial differential equations and Laplacians in Euclidean spaces together with the Lebesgue measure and its counterpart on manifolds have played a central role in understanding natural phenomena. In many areas, calculus is obstructed as in singular spaces, computer vision, learning theory, and quantum field theory. In vision it would be useful to do analysis on the space of images and an image is a function on a patch.

The point of view taken in this talk is to benefit from the Hodge theory to develop pattern analysis on general probability spaces without Lebesgue structure. This starts with a set X equipped with a distance d (which yields analysis like PDE and heat equations) as well a probability measure ρ (measuring the distribution of objects like images in X).

Let $\ell \in \mathbb{Z}_+$. The space $L^2(X^{\ell+1}) = L^2_\rho(X^{\ell+1})$ consists of ℓ-forms. The Hodge operator or co-boundary $\delta : L^2(X^{\ell+1}) \to L^2(X^{\ell+2})$ is defined by

$$\delta f(x^0, \dots, x^{\ell+1}) = \sum_{i=0}^{\ell+1} (-1)^i f(x^0, \dots, \widehat{x^i}, \dots, x^{\ell+1}).$$

Its dual $\delta^* = \partial : L^2(X^{\ell+2}) \to L^2(X^{\ell+1})$ is called the boundary operator.

The Laplacian on ℓ-forms is defined to be the operator $\Delta : L^2(X^{\ell+1}) \to L^2(X^{\ell+1})$ given by $\Delta = \delta\partial + \partial\delta$. If we denote Harm to be the space of all harmonic functions in $L^2(X^{\ell+1})$ satisfying $\Delta f = 0$, then we have the following Hodge decomposition (L^2 theory) [1].

Theorem 1. $L^2(X^{\ell+1}) = \mathrm{Im}\partial + \mathrm{Im}\delta + \mathrm{Harm}$

The Hodge operator δ can be generalized to a weighted setting with a symmetric and positive function K on $X \times X$. To see this, let $A_{\ell+1}$ be the weight function on $X^{\ell+1}$ given by $A_{\ell+1}(x^0, \dots, x^\ell) = \Pi_{i \neq j} \left(K(x^i, x^j)\right)^{1/2}$ for $\ell \geq 1$ while $A_1 \equiv 1$. Then the Hodge operator $\delta = \delta_K$ is from the weighted space $L^2_{\rho A_{\ell+1}}(X^{\ell+1})$ to the weighted space $L^2_{\rho A_{\ell+2}}(X^{\ell+2})$. Its dual $\delta^* = \partial : L^2_{\rho A_{\ell+2}}(X^{\ell+2}) \to L^2_{\rho A_{\ell+1}}(X^{\ell+1})$ is given by

$$\partial f(x^0, \dots, x^\ell) = \sum_{i=0}^{\ell+1} (-1)^i \int_X f(x^0, \dots, x^{i-1}, u, x^i, \dots, x^\ell) \Pi_{j=0}^{\ell} K(x^j, u) d\rho(u).$$

The Hodge operator and induced Laplacian can be used for learning theory. Consider the case $\ell = 0$ in the weighted setting with K being a Mercer kernel on

X. Then $A_1 \equiv 1$ and $A_2(x^0, x^1) = K(x^0, x^1)$. The Laplacian $\Delta = \partial\delta : L^2_\rho(X) \to L^2_\rho(X)$ on 0-forms takes the form

$$\Delta f(x) = 2D(x)f(x) - 2L_K f(x),$$

where $D(x) = \int_X K(x,u)d\rho(u)$ and L_K is the integral operator on $L^2_\rho(X)$ or the reproducing kernel Hilbert space $\mathcal{H}_K$ given by $L_K f(x) = \int_X K(x,u)f(u)d\rho(u)$.

The operator Δ can also be considered as one on $\mathcal{H}_K$. It can be discretized by a sample $\{x_i\}_{i=1}^m$ drawn from ρ. The function $D \in \mathcal{H}_K$ can be discretized as $\frac{1}{m}\sum_{i=1}^m K_{x_i}$ where $K_x = K(\cdot, x) \in \mathcal{H}_K$. The operator $L_K : \mathcal{H}_K \to \mathcal{H}_K$ can be approximated by a finite-rank one $\frac{1}{m}S_{\mathbf{x}}^T S_{\mathbf{x}}$ (induced by a sample operator $S_{\mathbf{x}}$ as in [2]) defined as $\frac{1}{m}S_{\mathbf{x}}^T S_{\mathbf{x}} f = \frac{1}{m}\sum_{i=1}^m \langle \cdot, K_{x_i}\rangle_K K_{x_i}$.

Theorem 2. Assume $\kappa := \sqrt{\sup_{x\in X} K(x,x)} < \infty$. With confidence $1-\delta$,

$$\left\| \frac{1}{m} S_{\mathbf{x}}^T S_{\mathbf{x}} - L_K \right\|_{\mathcal{H}_K \to \mathcal{H}_K} \leq \frac{4\kappa^2 \log\big(2/\delta\big)}{\sqrt{m}}.$$

Consider another weighted setting (corresponding to adjacency matrix of a graph X). Let $\alpha > 0$ and a subset of $X^{\ell+1}$ given by $\mathcal{U}_\alpha^{\ell+1} = \{(x^0, \ldots, x^\ell) \in X^{\ell+1} : d(x^i, p) \leq \alpha$ for some $p \in X$, and all $i\}$ (it equals $X^{\ell+1}$ when α is large enough). The Hodge operator $\delta = \delta_\alpha$ can be regarded as one from $L^2_\rho(\mathcal{U}_\alpha^{\ell+1})$ to $L^2_\rho(\mathcal{U}_\alpha^{\ell+2})$. Its dual $\partial : L^2_\rho(\mathcal{U}_\alpha^{\ell+2}) \to L^2_\rho(\mathcal{U}_\alpha^{\ell+1})$ is given by

$$\partial f(x^0, \ldots, x^\ell) = \sum_{i=0}^{\ell+1} (-1)^i \int_{S_{x^0,\cdots,x^\ell}} f(x^0, \ldots, x^{i-1}, u, x^i, \ldots, x^\ell) d\rho(u).$$

Here $S_{x^0,\cdots,x^\ell}$ denotes the slice $\{t \in X : (x^0, \ldots, x^\ell, t) \in \mathcal{U}_\alpha^{\ell+2}\}$. In this setting we have the following Hodge decomposition [1] where the space Harm of harmonic functions is defined by the corresponding Laplacian.

Theorem 3. For any $\alpha > 0$ and $\ell \in \mathbb{Z}_+$, we have

$$L^2_\rho(\mathcal{U}_\alpha^{\ell+1}) = \mathrm{Im}\partial + \mathrm{Im}\delta + \mathrm{Harm}.$$

The space of harmonic functions and in general eigenfunctions of the above Laplacian would lead to some applications in pattern analysis [3] as the graph Laplacian [4] does.

References

[1] N. Smale and S. Smale, *A general Hodge theory*, in preparation.

[2] S. Smale and D. X. Zhou, *Learning theory estimates via integral operators and their approximations*, Constr. Approx. **26** (2007), 153–172.

[3] S. Smale and D. X. Zhou, *Geometry on probability spaces*, preprint, 2008.

[4] M. Belkin and P. Niyogi, *Laplacian eigenmaps for dimensionality reduction and data representation*, Neural Comput. **15** (2003), 1373–1396.