Diffusion equations:
from Euclidean space to Graphs

MATH 697 AM:ST

November 7th, 2017

Warm up

Convolutions and differentiability

Consider the following:

A function $K : \mathbb{R}^d \mapsto \mathbb{R}$ which is infinitely differentiable, with

$$K(y) \geq 0 \; \forall \; y, \;\; K(y) = 0 \text{ if } |y| \geq 1, \;\; \int_{\mathbb{R}^d} K(y) \, dy = 1,$$

and which is spherically symmetric, that is

$$K(y) = K(y') \text{ if } |y| = |y'|.$$

For $\delta > 0$, define

$$K_\delta(y) = \delta^{-d} K(\delta^{-1} y).$$

Warm up

Convolutions and differentiability

Consider a domain $D \subset \mathbb{R}^d$, and $f : D \mapsto \mathbb{R}$ an integrable function,

$$\int_D |f(x)|\, dx < \infty$$

For $\delta > 0$ define the domain

$$D_\delta = \{x \ \mid d(x, \partial D) > \delta\}$$

Warm up

Convolutions and differentiability

Lemma

For $x \in D_\delta$, *define*

$$f_\delta(x) = (f * K_\delta)(x)$$

Then, f_δ *is infinitely differentiable in* D_δ, *and*

$$\partial^\alpha f_\delta(x) = (f * \partial^\alpha K_\delta)(x)$$

Minima of Functionals
RECAP

Hilbert's 19th Problem (1900)

Show that the minima for the functional

$$\mathcal{J}(f) = \int_D L(\nabla f, f, x)\, dx$$

are always analytic functions of x (under certain specific conditions we will not specify).

Minima of Functionals

RECAP

Hilbert's 19th Problem (1900)

1. Schauder (1930's): If there exists a $C^{1,\alpha}$ minimizer, then this minimizer must be analytic.
2. De Giorgi (1957), Nash (1958) showed that there were $C^{1,\alpha}$ minimizers. Their method relied greatly on new regularity estimates for partial differential equations.

Minima of Functionals

RECAP

The Dirichlet Energy

Given a bounded domain $D \subset \mathbb{R}^d$ with smooth boundary, and a continuous function

$$\phi : \partial D \mapsto \mathbb{R}$$

Problem

Find a function $f : D \mapsto \mathbb{R}$ equal to ϕ on ∂D, minimizing

$$\mathcal{J}(f) = \frac{1}{2} \int_D |\nabla f|^2 \, dx$$

Minima of Functionals

RECAP

The Dirichlet Energy

Theorem

If ϕ is differentiable, there is a unique continuous function

$$f : \overline{D} \mapsto \mathbb{R},\ f = \phi \text{ on } \partial D,$$

which is infinitely differentiable in the interior of D, and

$$\Delta f = 0.$$

Mean Value Property

RECAP

The MVP

Theorem

Let $f : \overline{D} \mapsto \mathbb{R}$ be a continouous function which is twice differentiable and harmonic in its interior.

Then, if x_0 is an interior point of D and r is strictly smaller than the distance from x_0 to ∂D, we have that

$$(\text{Average of } f \text{ over } \partial B_r(x_0)) = f(x_0).$$

Harmonic Functions and the MVP

Higher differentiability of harmonic functions

Suppose we are given a continuous function $f : \overline{D} \mapsto \mathbb{R}$ which is twice differentiable and harmonic in D.

Lemma

If $x \in D$ is a distance larger than δ from ∂D, then

$$f(x) = (f * K_\delta)(x)$$

Harmonic Functions and the MVP

Higher differentiability of harmonic functions

Theorem

A continuous function $f : \overline{D} \mapsto \mathbb{R}$ *which is twice differentiable and harmonic in* D *is always infinitely differentiable.*

Furthermore, for any index α *there is a constant* $C(d, \alpha)$ *such that in* D_δ *we have the estimate*

$$|\partial^\alpha f(x)| \leq C(d, \alpha)\delta^{-|\alpha|}\|f\|_{L^1(D)}.$$

Harmonic Functions and the MVP

Stability properties of Harmonic Functions

An important consequence of the higher differentiability and the estimate above, is the following stability property for harmonic functions (hinted at last time).

If a sequence of harmonic functions converge (locally) uniformly to a function, then this function is itself harmonic and the derivatives also converge (locally) uniformly.

The Fractional Laplacian

Last time, we defined the fractional Laplacian of order $\alpha \in [0, 2]$, by the formula

$$L_\alpha f(x) := C(d, \alpha) \int_{\mathbb{R}^d} \frac{f(y) - f(x)}{|y - x|^{d+\alpha}} dy$$

the constant $C(d, \alpha)$ is explicitly defined but we will not concern ourselves with its exact form.

The Fractional Laplacian

Last time, we defined the fractional Laplacian of order $\alpha \in [0, 2]$, by the formula

$$L_\alpha f(x) := C(d, \alpha) \int_{\mathbb{R}^d} \frac{f(y) - f(x)}{|y - x|^{d+\alpha}} dy$$

the constant $C(d, \alpha)$ is explicitly defined but we will not concern ourselves with its exact form.

We observed that $(-L_1) = (-\Delta)^{\frac{1}{2}}$, and that in general

$$-L_\alpha = (-\Delta)^{\frac{\alpha}{2}}$$

The Fractional Laplacian

The Nonlocal Dirichlet Energy: given $f : \mathbb{R}^d \mapsto \mathbb{R}$

$$J_\alpha(f) = \frac{1}{2} C(d,\alpha) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|f(x) - f(y)|^2}{|x-y|^{d+\alpha}} \, dxdy$$

Minimization of J_α leads to the equation $L_\alpha f = 0$, just as for the case of the Laplacian and the L^2 norm of $|\nabla f|$.

Graphs

Vertices, Edges, and weights

A (finite) weighted graph G is most typically, described as

$$G = (V, E, w)$$

$$\begin{aligned} V = &\ \text{a (finite) set, the set of } \textit{vertices} \\ E = &\ \text{a subset of } V \times V, \text{ the set of } \textit{edges} \\ w : E \mapsto \mathbb{R} &\ \text{the } \textit{weight} \text{ function} \end{aligned}$$

It is said that $x \sim y$ if $(x, y) \in E$,

Graphs

Vertices, Edges, and weights

Some simplification

It is usually preferable to think simply of the graph as being a set G (forget about distinguishing V), coupled with a (non-negative) weight function w,

$$w : G \times G \mapsto \mathbb{R}$$

with $w(x, y)$ denoted sometimes as w_{xy} for any $x, y \in G$.

One can think as E being the set of (x, y)'s such that $w_{xy} > 0$.

1. Subsets of $\mathbb{R}^p$.

Let $V = \{x_1, \ldots, x_N\}$ be a subset of $\mathbb{R}^p$, let

$$w(x_i, x_j) = h(x_i - x_j)$$

A popular choice is for h to be spherically symmetric, e.g.

$$\frac{1}{Z_\sigma} e^{-\frac{|x_i - x_j|^2}{\sigma}}, \quad \chi_{B_\sigma(0)}(x_i - x_j)$$

where $\sigma > 0$ is a tuning parameter.

2. Subsets of a metric space.

Let $V = \{x_1, \ldots, x_N\}$ be a subset of (X, ρ), a metric space. Then, define

$$w(x_i, x_j) = h\left(\frac{\rho(x_i, x_j)^2}{\sigma^2}\right)$$

Popular selections for h are

$$h(t) = \sqrt{t}$$
$$h(t) = e^{-t}$$
$$h(t) = \chi_{[0,1]}(t)$$

Graphs

Setup

Given a vertex x, it's **degree** is the number

$$d(x) = \sum_{y \in V} w_{xy}$$

The **normalized weight function** is then defined by

$$K(x, y) := \frac{1}{d(x)} w_{xy}$$

so that for fixed x, $K(x, \cdot)$ is a probability distribution over G,

$$\sum_{y \in V} K(x, y) = 1.$$

The Dirichlet Problem on graphs and it's use in Semi-supervised learning

MATH 697 AM:ST

November 9th, 2017

Warm up

A **path** in a graph G is a sequence of points x_i $(i = 1, \ldots, m)$ such that

$$x_i \sim x_{i+1} \text{ for } i = 1, 2, \ldots, m - 1$$

Two points $x, y \in G$ are said to be connected if there exists a path with $x_1 = x$ and $x_m = y$.

A subset D of G is said to be **connected** if given two points in D then they can be connected by a path exclusively made out of points in D.

Warm up

We consider the Laplacian of a function $f : G \mapsto \mathbb{R}$

$$\Delta f(x) = \sum_{y \in G} w_{xy}(f(y) - f(x)).$$

Then, we observe the following:

$$f(x_0) \geq f(x) \;\; \forall\, x \in G \Rightarrow \Delta f(x_0) \leq 0.$$

Moreover, if $f(x_0) > f(x_1)$ for at least one $x_1 \sim x_0$, then

$$\Delta f(x_0) < 0.$$

Graph Laplacians

RECAP

We define various operators all with equal claims to be called a Laplacian.

First, the **Combinatorial Laplacian** (or just **the Laplacian**)

$$\Delta f(x) = \sum_{y \in G} (f(y) - f(x)) w_{xy}$$

Graph Laplacians

RECAP

We define various operators all with equal claims to be called a Laplacian.

Secondly, we have the **Random Walk Laplacian**

$$\Delta f(x) = \sum_{y \in G} (f(y) - f(x)) K(x, y)$$

Graph Laplacians
RECAP

We define various operators all with equal claims to be called a Laplacian.

Last but not least, we have the **Symmetric Laplacian**

$$\Delta f(x) = \frac{1}{\sqrt{d(x)}} \sum_{y \in G} w_{xy} \left(\frac{g(y)}{\sqrt{d(x)}} - \frac{g(x)}{\sqrt{d(x)}} \right)$$

Graph Laplacians

RECAP

The **Combinatorial Laplacian** (or just **the Laplacian**)

$$\Delta f(x) = \sum_{y \in G} (f(y) - f(x)) w_{xy}$$

The **Random Walk Laplacian**

$$\Delta f(x) = \sum_{y \in G} (f(y) - f(x)) K(x, y)$$

The **Symmetric Laplacian**

$$\Delta f(x) = \frac{1}{\sqrt{d(x)}} \sum_{y \in G} w_{xy} \left(\frac{g(y)}{\sqrt{d(x)}} - \frac{g(x)}{\sqrt{d(x)}} \right)$$

The (strong) Maximum Principle

Consider a function $f : G \mapsto \mathbb{R}$ such that

$$\Delta f(x) = \sum_{y \in G} w_{xy}(f(y) - f(x)) = 0 \;\; \forall\, x \in G.$$

Let $M = \max_{x \in G} f(x)$, and let $x_0 \in G$ be such that

$$f(x_0) = M$$

Claim:

If x is connected to x_0 then we have $f(x_0) = M$

The (strong) Maximum Principle

Theorem

Let G be a connected graph, then the only functions $f : G \mapsto \mathbb{R}$ which are harmonic in G are the constants.

In other words: for a connected graph, the kernel of the linear map Δ is one dimensional and given by the constant function.

To have any interesting harmonic functions, we must ask they be harmonic *only in some portion* of G.

The Comparison Principle

Theorem

Let G be a graph, and $D \subset G$ a connected subset of the graph. Then, if $f_1, f_2 : G \mapsto \mathbb{R}$ are such that

$$\Delta f_1 = \Delta f_2 = 0 \text{ in } D, \quad \text{and } f_1 \leq f_2 \text{ in } G \setminus D$$

Then, we have

$$f_1 \leq f_2 \text{ in } D$$

Measuring Smoothness

Let $f : G \mapsto \mathbb{R}$, and define the norms

$$\|f\|_{L^2} := \left(\sum_{x \in G} |f(x)|^2 \right)^{\frac{1}{2}}$$

$$\|f\|_{H_w} := \left(\sum_{x \in G} |f(x)|^2 + \sum_{x,y \in G} |f(x) - f(y)|^2 \omega_{xy} \right)^{\frac{1}{2}}$$

as well as

$$\|f\|_{\dot{H}_w} := \left(\sum_{x,y \in G} |f(x) - f(y)|^2 \omega_{xy} \right)^{\frac{1}{2}}$$

Measuring Smoothness

While $\|f\|_{L^2}$ measures simply the average “size” of f, $\|f\|_{\dot{H}_w}$ measures the average size of its oscillations.

Observe that if $\|f\|_{\dot{H}_w} = 0$, then f must be constant in each connected component of the graph.

Thus, the larger $\|f\|_{\dot{H}_w}$ is with respect to $\|f\|_{L^2}$, the more the function f is oscillating with respect to its size.

The Laplacian is a symmetric operator

We introduce an inner product for functions in G,

$$\langle f, g \rangle = \sum_{x \in G} f(x) g(x)$$

Henceforth assume that G has symmetric weights, that is

$$w_{xy} = w_{yx}$$

The Laplacian is a symmetric operator

Let Δ be the combinatorial Laplacian, and observe that

$$\sum_{x \in G} \Delta f(x) g(x) = \sum_{x \in G} \left(\sum_{y \in G} (f(y) - f(x)) w_{xy} \right) g(x)$$

The Laplacian is a symmetric operator

Then, expanding this sum, we have

$$\langle \Delta f, g \rangle = \sum_{x,y \in G} f(y) g(x) w_{xy} - \sum_{x \in G} f(x) g(x) d(x)$$

where $d(x) = \sum_{y \in G} w_{xy}$. Since $w_{xy} = w_{yx}$, it follows that

$$\langle \Delta f, g \rangle = \langle f, \Delta g \rangle$$

Therefore, Δ is symmetric with respect to this inner product.

The Laplacian is a symmetric operator

Positivity

One further expression, which is reminiscent of the Green identity, is the following (to be proved later)

$$-\langle \Delta f, g \rangle = -\frac{1}{2} \sum_{x,y \in G} w_{xy}(f(y) - f(x))(g(y) - g(x))$$

It follows that for any $f : G \mapsto \mathbb{R}$,

$$-\langle \Delta f, f \rangle \geq 0$$

The Laplacian is a symmetric operator

Positivity

Furthermore, if G is connected, then

$$-\langle \Delta f, f \rangle = 0$$

if and only if f is equal to a **constant**.

With this knowledge in hand, we obtain a good picture of the eigenfunction decomposition associated to Δ.

Eigenfunctions of Δ

There is a family of functions

$$\{\phi_n\}_{n=0}^N, \quad \phi_n : G \mapsto \mathbb{R}$$

as well as numbers $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$, such that

$$-\Delta\phi_n(x) = \lambda_n\phi_n(x) \quad \forall\, x \in G.$$

Moreover, the ϕ_n are orthonormal,

$$\langle \phi_n, \phi_m \rangle = \delta_{nm}.$$

Eigenfunctions of Δ

Decrasing Smoothness of ϕ_n

Observe that

$$\|\phi_n\|^1_{\dot{H}_w} = \sum_{x,y \in G} w_{xy}(\phi_n(y) - \phi_n(x))^2$$
$$= -2\langle \Delta\phi_n, \phi_n \rangle$$

Therefore,

$$\|\phi_n\|^1_{\dot{H}_w} = 2\lambda_n \|\phi_n\|^2_{L^2}$$

As the λ_n is increasing with n, we see that the ϕ_n become increasingly more oscillatory as n gets larger and larger.

The Heat Equation

We consider the heat equation in all of G, with some initial datum $u_0 : G \mapsto \mathbb{R}$,

$$\begin{aligned} \dot{u} &= \Delta u \quad \text{in } G \times (0, \infty), \\ u &= u_0 \quad \text{at } t = 0. \end{aligned}$$

The Heat Equation

By the orthogonality of the ϕ_n, it follows that u_0 can be expressed as

$$u_0 = \sum_{n=0}^{N} \alpha_n \phi_n \quad \text{where } \alpha_n = \langle \phi_n, u_0 \rangle.$$

Therefore, for $t > 0$, the solution to the heat equation is

$$u(t) = \sum_{n=0}^{N} \alpha_n e^{-t\lambda_n} \phi_n$$

The Heat Equation

By the orthogonality of the ϕ_n, it follows that u_0 can be expressed as

$$u_0 = \sum_{n=0}^{N} \alpha_n \phi_n \quad \text{where } \alpha_n = \langle \phi_n, u_0 \rangle.$$

Therefore, for $t > 0$, the solution to the heat equation is

$$u(t) = \sum_{n=0}^{N} e^{-t\lambda_n} \langle \phi_n, u_0 \rangle \phi_n$$

The Heat Equation

By the orthogonality of the ϕ_n, it follows that u_0 can be expressed as

$$u_0 = \sum_{n=0}^{N} \alpha_n \phi_n \quad \text{where } \alpha_n = \langle \phi_n, u_0 \rangle.$$

Therefore, for $t > 0$, the solution to the heat equation is

$$u(x,t) = \sum_{n=0}^{N} \sum_{y \in G} e^{-t\lambda_n} \phi_n(y) u_0(y) \phi_n(x)$$

The Heat Equation

This leads to an analogue of the heat kernel $H(t, x, y)$, namely, a function such that

$$u(x,t) = \sum_{y \in G} H(t,x,y) u_0(y)$$

This $H(t, x, y)$ is given via the eigenfunction decomposition

$$H(t,x,y) = \sum_{n=0}^{N} e^{-t\lambda_n} \phi_n(x) \phi_n(y)$$

The Dirichlet Problem Revisited

Fix some subset $D \subset G$, and suppose we are given a function

$$g : G \setminus D \mapsto \mathbb{R}$$

Then, we aim to find $f : G \mapsto \mathbb{R}$, such that

$$f = g \text{ in } G \setminus D,$$

and with f minimizing the discrete Dirichlet energy

$$J(f) := \frac{1}{2} \sum_{x,y \in G} w_{xy}(f(x) - f(y))^2$$

The Dirichlet Problem Revisited

Let $f_s = f + s\phi$, then

$$\frac{d}{ds}J(f_s) = \frac{d}{ds}\left\{\frac{1}{2}\sum_{x,y\in G} w_{xy}(f_s(x) - f_s(y)^2\right\}$$
$$= \sum_{x,y\in G} w_{xy}(f_s(x) - f_s(y)(\phi(x) - \phi(y))$$

At $s = 0$, this results in

$$\sum_{x,y\in G} w_{xy}(f(x) - f(y)(\phi(x) - \phi(y))$$

The Dirichlet Problem Revisited

$$\sum_{x,y \in G} w_{xy}(f(x) - f(y)(\phi(x) - \phi(y))$$

$$= \left\{ \sum_{x,y \in G} w_{xy}(f(x) - f(y))\phi(x) \right\} - \left\{ \sum_{x,y \in G} w_{xy}(f(x) - f(y))\phi(y) \right\}$$

Since $w_{xy} = w_{yx}$ means that

$$\sum_{x,y \in G} w_{xy}(f(x) - f(y))\phi(y) = \sum_{x,y \in G} w_{xy}(f(y) - f(x))\phi(x)$$

The Dirichlet Problem Revisited

$$\sum_{x,y\in G} w_{xy}(f(x)-f(y)(\phi(x)-\phi(y))$$

$$= 2\sum_{x,y\in G} w_{xy}(f(x)-f(y))\phi(x)$$

$$= -2\sum_{x\in G}\phi(x)\sum_{y\in G} w_{xy}(f(y)-f(x))$$

$$= -2\sum_{x\in G}\phi(x)\Delta f(x)$$

The Dirichlet Problem Revisited

In conclusion,

$$\frac{d}{ds}_{|s=0} J(f_s) = -2 \sum_{x \in G} \phi(x) \Delta f(x)$$

In particular,

$$\sum_{x \in G} \phi(x) \Delta f(x) = 0$$

for any function $\phi : G \mapsto \mathbb{R}$ with $\phi(x) = 0$ if $x \in G \setminus D$.

The Dirichlet Problem Revisited

There is a unique minimizer f, and it is characterized by

$$\begin{cases} Lf(x) &= 0 \qquad \text{if } x \in D, \\ f(x) &= g(x) \text{ if } x \in G \setminus D. \end{cases}$$

The Dirichlet Problem Revisited

There is a unique minimizer f, and it is characterized by

$$\begin{cases} Lf(x) & = 0 \qquad \text{if } x \in D, \\ f(x) & = g(x) \text{ if } x \in G \setminus D. \end{cases}$$

Compare with the Dirichlet problem for the Laplacian

$$\min \int_D |\nabla f(x)|^2 \, dx, \;\; f = g \text{ on } \partial D.$$

and the one for the fractional Laplacian ($\alpha \in (0,2)$),

$$\min \int_{\mathbb{R}^{2d}} (f(x) - f(y)|^2 |x-y|^{-d-2\alpha} \, dx, \;\; f = g \text{ in } \mathbb{R}^d \setminus D.$$

The Dirichlet Problem Revisited

Thanks to the variational characterization of harmonic functions and the comparison theorem, we have the following:

Theorem

Let $D \subset G$ be a connected subset of G. Then, given $g : G \setminus D \mapsto \mathbb{R}$ there exists one, and exactly one, function f which solves

$$\begin{cases} Lf(x) &= 0 \qquad \text{if } x \in D, \\ f(x) &= g(x) \text{ if } x \in G \setminus D. \end{cases}$$

The Dirichlet Problem Revisited

The solution to

$$\begin{cases} Lf(x) &= 0 \qquad \text{if } x \in D, \\ f(x) &= g(x) \text{ if } x \in G \setminus D. \end{cases}$$

Is given by

$$f(x) = \sum_{y \in G \setminus D} g(y) P(x, y)$$

for a certain kernel $P(x, y)$ computable from the w_{xy}.