Diffusion equations: from Euclidean space to Graphs

MATH 697 AM:ST

October 31st, 2017

Warmup

The Euler-Lagrange Equation

One of the central objects in mathematics and physics is the class of functionals of the form

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

Defined for differentiable functions f defined in D.

The function $L(p)$ ($p \in \mathbb{R}^d$) is known as the *Lagrangian*.

Warmup

The Euler-Lagrange Equation

Consider $f_0 : D \mapsto \mathbb{R}$, a twice differentiable function such that

$$\mathcal{J}(f_0) \leq \mathcal{J}(f)$$

for any other differentiable function of the form $f = f_0 + \phi$, where ϕ is twice differentiable and has compact support in D (i.e. $\phi \equiv 0$ in a neighborhood of ∂D, if D is bounded).

Theorem (Euler-Lagrange)

The function f_0 solves the (partial) differential equation

$$\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(\left(\frac{\partial L}{\partial p_i} \right) (\nabla f_0) \right) = 0$$

Warmup

The Euler-Lagrange Equation

More general equations are obtained if one considers Lagrangians with extra dependence on f, such as

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

where $L(p, z, x)$ is smooth in all its variables. In this case, the Euler-Lagrange equation is

$$\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial p_i}(\nabla f_0, f_0, x) \right) = \frac{\partial L}{\partial z}(\nabla f_0, f_0, x)$$

Example: Dirichlet's Principle

Consider the functional (*Dirichlet energy*)

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

It was Dirichlet who observed that minimization of this functional leads to harmonic functions. Indeed, note that

$$\begin{aligned} \frac{d}{ds}_{|s=0}\mathcal{E}_D(f_s) &= \int_D \nabla f \cdot \nabla \dot{f} \, dx \\ &= \int_D (-\Delta f)\dot{f} \, dx \end{aligned}$$

The last identity holding as long as $\dot{f}$ vanishes on ∂D.

Last time

1. Some background on Fourier analysis and the heat equation.
2. The smoothing effect of the heat equation: examples.
3. Eigenfunctions of the Laplacian.

This week

1. The smoothing effect of the heat equation in $\mathbb{R}^d$.
2. The mean value property for harmonic functions.
3. The fractional Laplacian.
4. Graph Laplacians and discrete diffusions.
5. Semi-supervised learning via harmonic functions.

The heat equation in $\mathbb{R}^d$

Let us revisit the heat equation, this time in the whole space.

The Cauchy problem for the heat equation in $\mathbb{R}^d$ requires finding, for a given $f_0(x)$, the unique function $f(x,t)$ solving

$$\begin{cases} \partial_t f &= \Delta f \text{ in } \mathbb{R}^d \times (0,\infty) \\ f &= f_0 \quad \text{at } t = 0 \end{cases}$$

The heat equation in $\mathbb{R}^d$

The unique solution to the heat equation is given by taking the convolution of f_0 with respect to rescaled Gaussians, that is

$$f(x,t) = (f_0 * \Gamma_t)(x)$$

where, for every t, we have

$$\Gamma_t(x) = t^{-\frac{d}{2}}\Gamma(t^{-\frac{1}{2}}x)$$

and

$$\Gamma(x) = \frac{1}{(4\pi)^{\frac{d}{2}}}e^{-\frac{|x|^2}{4}}$$

The heat equation in $\mathbb{R}^d$

Connection with stochastic processes

Let us write the integral in full, we have

$$f(x,t) = \int_{\mathbb{R}^d} f_0(y) \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-\frac{|x-y|^2}{4t}} \, dy$$

Observe that this can be thought of in terms of an *expectation*,

$$f(x,t) = \mathbb{E}[f_0(x+B_t)]$$

where B_t is Brownian motion, this is related to the famous *Feynman-Kac formula*.

The heat equation in $\mathbb{R}^d$

As we saw in terms of Fourier series, the heat equation has a strong smoothing effect. This is made manifest for the equation in the whole space by the formula we just derived.

Indeed, note that if $t > 0$, then

$$\nabla f(x,t) = f_0 * \nabla \Gamma_t,$$

and more generally,

$$D^k f(x,t) = f_0 * D^k \Gamma_t.$$

The heat equation in $\mathbb{R}^d$

For $t > 0$, derivatives of Γ_t of given order k are bounded in $\mathbb{R}^d$

Thus, we have the following estimate for the derivatives of f,

$$\sup_{\mathbb{R}^d} |\tfrac{\partial^\alpha}{\partial x^\alpha} f(x,t)| \leq \frac{C(d,\alpha)}{t^{\frac{d}{2}+k}} \int_{\mathbb{R}^d} |f_0(y)|\, dy$$

and we conclude that if the intial data f_0 is integrable, then the solution $f(x,t)$ will be infinitely differentiable for positive t.

The heat equation in $\mathbb{R}^d$

Gradient Flow Structure

The heat equation has another interpretation: it is a gradient flow for the Dirichlet energy.

This is hinted at by the following computation, let

$$\mathcal{E}(f) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla f|^2 \, dx$$

Then, if f solves the heat equation,

$$\begin{aligned} \frac{d}{dt}\mathcal{E}(f) &= \int_{\mathbb{R}^d} \nabla f \cdot \nabla f_t \, dx \\ &= -\int_{\mathbb{R}^d} (\Delta f) f_t \, dx = -\int_{\mathbb{R}^d} (\Delta f)^2 \, dx, \end{aligned}$$

as it turns out, the last term can be understood as $-|\nabla_f \mathcal{E}|^2$!

The same computation can be done for the problem in a bounded domain with various boundary conditions, i.e.

$$\begin{aligned} \partial_t f = \Delta f \ \text{ in } D \times (0, \infty) \\ f = 0 \quad \text{on } (\partial D) \times (0, \infty) \\ f = f_0 \quad \text{at } t = 0 \end{aligned}$$

Once again, we can show that

$$\frac{d}{dt} \int_D |\nabla f|^2 \, dx = - \int_{\mathbb{R}^d} (\Delta f)^2 \, dx$$

Then, by letting $t \to \infty$, we expect the heat equation to flow towards the minima of the functional

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

over all functions with zero boundary values on ∂D.
In this case, there is one minimizer, and it is given by the constant zero function, but this may not be the case in general!.

Functionals of the form

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

where $L(p, z, x)$ is convex with respect to the $p \in \mathbb{R}^d$ variable, have been of great importance (observe the Dirichlet energy is one such example), another example is given by the *area functional*

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

Diffusion equations: from Euclidean space to Graphs (continued)

MATH 697 AM:ST

November 2nd, 2017

Warm up

The Comparison Property of the Laplacian

Observation:

Let $f, g : D \mapsto \mathbb{R}$ be twice differentiable functions.

Suppose that $f \leq g$ everywhere in D, and that at some point x_0

$$f(x_0) = g(x_0).$$

Then, we have

$$\Delta f(x_0) \leq \Delta g(x_0).$$

Warm up

The Comparison Property of the Laplacian

Observation: (continued)

Suppose that f is twice differentiable in D, continuous in $\overline{D}$, and such that

$$\Delta f(x) < 0 \ \forall \, x \in D.$$

Then, we have that

$$\max_{D} f = \max_{\partial D} f.$$

Warm up

The Comparison Property of the Laplacian

Theorem (The Comparison Principle)

Let $f, g : \overline{D} \mapsto \mathbb{R}$ be two continuous functions which are twice differentiable in D. Suppose that $f \leq g$ in ∂D, and that

$$\Delta f \geq \Delta g \text{ in } D.$$

Then, we have

$$f \leq g \text{ in } D.$$

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

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

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

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

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

Mean Value Property

The MVP

Consequences of the Mean Value Property

1. Average over $B_r(x_0)$, not just $\partial B_r(x_0)$, is also $f(x_0)$.
2. A harmonic function cannot achieve its maximum at an interior point, unless it is constant (*Maximum Principle*).
3. A harmonic function is infinitely differentiable in the interior of D.
4. If a sequence of harmonic functions converges (locally) uniformly to a function, then this function is itself harmonic and the sequence of respective derivatives all converge (locally) uniformly to the derivatives of the function.

The Fractional Laplacian

Let us now consider a different operator related to the Laplacian.

In fact, this is a family of operators L_α, indexed by $\alpha \in [0, 2]$. L_α will take α derivatives, not 2.

Just as Δf measures the *infinitesimal mean oscillation* at x, $L_\alpha f(x)$ will measure the mean oscillation at multiple scales, with the values of f at points near x having the most weight.

The Fractional Laplacian

The operator is defined as follows, for $\alpha \in [0,2]$.

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

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

The operator L_α is called the *Fractional Laplacian* of order α.

The Fractional Laplacian

The adjective fractional comes from the fact that $-L$ agrees with a fractional power of $-\Delta$. For instance, we have that

$$(-L_1)^2 f = L_1(L_1 f) = -\Delta f$$

so, it is fair to say that $L_1 = -(-\Delta)^{\frac{1}{2}}$.

Most importantly, if one fixes f, and one computes the Fourier transform of $L_\alpha f$, one has

$$\widehat{(L_\alpha f)}(y) = |y|^\alpha \hat{f}(y)$$

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 (non-negative) weight function w,

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

with $w(i, j)$ denoted as w_{ij} for any $i, j \in G$.

One can think as E being the set of (i, j)'s such that $w_{ij} > 0$.

Graphs

Example

Let $V = \{x_1, \dots, x_N\}$ be a subset of $\mathbb{R}^p$, let $E = V \times V$, and

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

Popular selections for h are

$$\frac{1}{Z_\sigma} e^{-\frac{|x_i - x_j|^2}{\sigma}}$$

$$\chi_{B_\sigma(0)}(x_i - x_j)$$

Graphs

Example

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

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

Popular selections for h are

$$h(t) = e^{-t^2}$$
$$h(t) = \chi_{[0,1]}(t)$$

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

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

From data sets to graphs

Local Similarities

If the data set amounts to points $\{x_1, \ldots, x_N\}$ in some metric space (X, ρ),

$$W_\sigma(x, y) = h\left(\frac{1}{\sigma^{\frac{1}{2}}}\rho(x, y)\right)$$

Graph Laplacians

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

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

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

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**

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