

Geometric Deep Learning Beyond Euclidean Domains

1 Geometric Priors

Definition 1.1. Our compact euclidean domain Ω

$$\Omega := \prod_{i \in I} [0, 1].$$

Definition 1.2. Classification

Let $x \in L^2 := L^2(\Omega)$ then $f : L^2 \rightarrow \mathcal{C}$ surjective is said to be a classification of L^2 on the set \mathcal{C} .

Definition 1.3. Training set

Let f be a classification of L^2 on \mathcal{C} and $\{x_i\}_{i \in I} \subset L^2$ then the set $\{(x_i, f(x_i))\}_{i \in I}$ is called a training set for f .

Proposition 1.1. The classification f is not injective

Let f be a classification of L^2 on \mathcal{C} then, given the inevitable noise acting on data, there exists a real positive ε such that $\forall (x, x_\varepsilon) \in L^2 \times L^2 : \int_{\Omega} |x - x_\varepsilon|^2 < \varepsilon$ we have that $f(x) = f(x_\varepsilon)$.

Given ideal data classification we can define two functions f -equivalent if and only if their images via the classification f are equal according to an equivalence on \mathcal{C} which so far can be any set.

Proposition 1.2. The relation \simeq is an equivalence relation

Let $x, y, z \in L^2$ we define $x \simeq y \iff f(x) = f(y)$ where f is a classification of L^2 on \mathcal{C} , then:

(i) $x \simeq x$

(ii) $x \simeq y \iff y \simeq x$

(iii) $x \simeq y, y \simeq z \implies x \simeq z$

Proof. (i),(ii) and (iii) follow from the the equivalence on \mathcal{C} by which they are defined. □

Definition 1.4. Translation operator

Let $x \in L^2$ and $v \in \Omega$ then $T_v : L^2 \rightarrow L^2$ such that $x(\xi) \mapsto x(\xi - v)$ is said to be a translation operator.

Definition 1.5. Local deformation operator

Let $x \in L^2$ and $\tau \in C^\infty(\Omega, \Omega)$ then $L_\tau : L^2 \rightarrow L^2$ such that $x(\xi) \mapsto x(\xi - \tau(\xi))$ is said to be a local deformation operator according to the smooth vector field τ .

Definition 1.6. Invariance

A classification f of L^2 on \mathcal{C} is said to be A -invariant, where $A : L^2 \rightarrow L^2$, if and only if $f(A(x)) = f(x) \forall x \in L^2$.

Definition 1.7. Equivariance

A classification f of L^2 on \mathcal{C} is said to be A -equivariant, where $A : L^2 \rightarrow L^2$, if and only if $f(A(x)) = A(f(x)) \forall x \in L^2$. This is well defined only if A is defined to act on \mathcal{C}

Proposition 1.3. If f is translation invariant then it is stable under local deformations

Let f be a translation invariant classification of L^2 on \mathcal{C} then $|f(L_\tau(x)) - f(x)| \approx |J_\tau|$ where $(J_\tau)_{ij} = (\frac{\partial \tau_i}{\partial \xi_j})$ under some misterious norm.

Proof. To be found... □

Aggiungere la struttura delle reti convoluzionali e la formula da minimizzare coi filtri di convoluzione

2 Graphs and Manifolds

Definition 2.1. Let \mathcal{G} be a graph where \mathcal{V} are its vertexes and \mathcal{E} are its edges, let $f, g \in L^2(\mathcal{V})$ and $F, G \in L^2(\mathcal{E})$ be real valued functions, we define $\langle f, g \rangle_{L^2(\mathcal{V})} := \sum_{\mathcal{V}} a_i f_i g_i$, $a_i \in \mathbb{R}$ and $\langle F, G \rangle_{L^2(\mathcal{E})} := \sum_{\mathcal{E}} w_{ij} F_{ij} G_{ij}$, $w_{ij} \in \mathbb{R}$. Let M be a manifold and TM its tangent bundle with a metric $\langle \cdot, \cdot \rangle_{TM} : TM^2 \rightarrow \mathbb{R}$, let $f, g \in L^2(M)$ and $F, G \in L^2(TM) := F : M \rightarrow TM$, given two scalar products $\langle f, g \rangle_{L^2(M)} := \int_M dx f g$ and $\langle F, G \rangle_{L^2(TM)} := \int_M dx \langle F, G \rangle_{TM}$.

Definition 2.2. Graph gradient and divergence

Let $f \in L^2(\mathcal{V})$ and $F \in L^2(\mathcal{E})$ we define $\text{grad} : L^2(\mathcal{V}) \rightarrow L^2(\mathcal{E})$ and $\text{div} : L^2(\mathcal{E}) \rightarrow L^2(\mathcal{V})$, such that $(\text{grad} f)_{ij} = f_i - f_j$ and $(\text{div} F)_i = \frac{1}{a_i} \sum_{j \in \mathcal{V} : (i,j) \in \mathcal{E}} w_{ij} F_{ij}$.

Proposition 2.1. Let $f \in L^2(\mathcal{V})$ and $F \in L^2(\mathcal{E}) : F_{ij} = -F_{ji}$ then $\langle f, \text{div} F \rangle_{L^2(\mathcal{V})} = \langle \text{grad} f, F \rangle_{L^2(\mathcal{E})}$, i.e. $\text{div}^\dagger = \text{grad}$.

Proof. $\sum_{\mathcal{V}} a_i f_i (\text{div} F)_i = \sum_{\mathcal{E}} w_{ij} F_{ij} (f_i - f_j) = \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V} : (i,j) \in \mathcal{E}} w_{ij} F_{ij} f_i$ thus $a_i (\text{div} F)_i = \sum_{j \in \mathcal{V} : (i,j) \in \mathcal{E}} w_{ij} F_{ij}$. \square

Definition 2.3. Manifold gradient and divergence

Let $f \in L^2(M)$ and $F \in L^2(TM)$ we define $(\text{grad} f)_i := \frac{\partial f}{\partial x_i}$ and $\text{div} F := \sum_i \frac{\partial F_i}{\partial x_i}$.

Proposition 2.2. Let $f \in L^2(M)$ and $F \in L^2(TM)$ then $\langle f, -\text{div} F \rangle_{L^2(M)} = \langle \text{grad} f, F \rangle_{L^2(TM)}$, i.e. $\text{div}^\dagger = \text{grad}$.

Proof. $\int_M dx \langle \text{grad} f, F \rangle_{L^2(TM)} = \int_M dx (\text{div}(fF) - f \text{div} F) = \int_{\partial M} f F + \int_M dx f (-\text{div} F)$, where some condition must be added to make the first integral vanish. \square

Definition 2.4. Path on a graph

We define a curve parametrization on a graph be some injective function $\gamma : I \subset \mathbb{Z} \rightarrow \mathcal{V}$, we shall call γ a curve meaning its image $\gamma(I)$, to satisfy the connectedness condition we want the simplicial complex $\{\gamma, [\gamma(n), \gamma(n+1)]_{n \in I}\}$ to have a 1 dimensional 0-homology group. We define path over a graph the sequence $\{[\gamma(n), \gamma(n+1)]_{n \in I} \subset \mathcal{E}$, we might as well call that γ because we are bad people.

Theorem 2.3. Gradient theorem on graphs

Let γ be a connected path on a graph, and $f \in L^2(\mathcal{V})$, than we have $\sum_{\gamma} (\text{grad} f) = D f_{\partial_0 \gamma}$, where we define $D f_{(i,j)} = f_i - f_j$, and by ∂_0 we mean the boundary operator.

Proof. Left to the reader... \square

In realtà se avessi la certezza che il bordo di un complesso simpliciale 1D connesso è vuoto o due punti starei meglio, esiste un teorema? boh

Theorem 2.4. Gauss theorem on graphs

Let $F \in L^2(\mathcal{E}) : F_{ij} = -F_{ji}$, let $\mathcal{A} \subset \mathcal{V}$ then if $a_i = w_{ij} = 1$ we have $\sum_{\mathcal{A}} (\text{div} F)_i = \sum_{\partial^0 \mathcal{A}} F_{ij}$.

Proof. First of all we recall $\partial^0 \mathcal{A} = \{(i,j) \in \mathcal{E}, i \in \mathcal{A}, j \in \mathcal{V} \setminus \mathcal{A}\}$, then we see that $\sum_{\mathcal{A}} (\text{div} F)_i = \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{V} : (i,j) \in \mathcal{E}} F_{ij} = \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{V} \setminus \mathcal{A} : (i,j) \in \mathcal{E}} F_{ij} + \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A} : (i,j) \in \mathcal{E}} F_{ij} = \sum_{\partial^0 \mathcal{A}} F_{ij} + \sum_{(i,j) \in \mathcal{A}^2} \text{adj}(\mathcal{A})_{ij} F_{ij}$ where since $\text{adj}(\mathcal{A})_{ij} = \text{adj}(\mathcal{A})_{ji}$ we have by renaming dummy indexes $\text{adj}(\mathcal{A})_{ij} F_{ij} = -\text{adj}(\mathcal{A})_{ij} F_{ji} = 0$. \square

Aggiungere la formalizzazione del cobordo e del bordo tramite gli operatori, ridefinizione della divergenza con il cobordo del vertice

Proposition 2.5. *The use of the coboundary operator makes sense only with antisymmetric functions on the edges, the antisymmetry of those function is somehow related to the orientation of surfaces.*

Proof. $\sum_{\partial^0} \sum_{i \in \mathcal{A}} i F_{ij} = \sum_{i \in \mathcal{A}} \sum_{\partial^0 i} F_{ij}$ if an edge is in the coboundary of two different vertexes of \mathcal{A} it will be count twice, that means zero times in \mathbb{Z}_2 , similarly for that same edge we would sum $F_{ij} + F_{ji} = 0$. \square

Definition 2.5. Graph laplacian

Let $f \in L^2(\mathcal{V})$ we have that $\langle \text{grad} f, \text{grad} f \rangle = \langle \text{div}(\text{grad} f), f \rangle =: \langle \Delta f, f \rangle = \langle f, \Delta f \rangle$, where $\Delta : L^2(\mathcal{V}) \rightarrow L^2(\mathcal{V})$ is the Laplacian.

Proposition 2.6. The laplacian represents the difference between the function and a local average of the function

(i) Graphs

Let $w_{ii} = 0$, and $\sum_j w_{ij} = a_i$, i.e. normalized laplacian, we have that $(\Delta f)_i = f_i \frac{\sum_j w_{ij}}{a_i} - \sum_j \frac{w_{ij}}{a_i} f_j$ which is a weighted average.

(ii) Manifolds (Let's just see it on an euclidean domain)

Let $f_0 \int_{\partial B} dx - \int_{\partial B} dx f(x) \simeq \int_{\partial B} dx \langle \text{grad} f, x \rangle = \int_B dx \Delta f$, where B is a ball centered in 0 (or at least has a boundary), we have that $f(0) - \frac{\int_{\partial B} dx f(x)}{\int_{\partial B} dx} \simeq \frac{\int_B dx \Delta f}{\int_{\partial B} dx}$. Gauss theorem could be used since the incremental vector x on a ball is parallel to $2x$ with is the gradient of the implicit function defining the ball.

Possibili approfondimenti interessanti:

(i) Studio di equazioni differenziali sui grafi con gli operatori sopra definiti

-Equazione del Calore

$$\frac{d(f_i)}{dt} = -c(\Delta f)_i$$

-Equazione di Schrödinger

$$\frac{d(f_i)}{dt} = -c(\Delta f)_i + U_i f_i$$

-Equazione di Navier-Stokes

-Equazione di continuità

-altre

(ii) Vincolare l'apprendimento di funzioni a divergenza nulla delle edge tramite H^0 , eventuale applicazione a flussi incompressibili

(iii) Definire il rotore per 2-simplessi e fare l'analogo con H_1 , eventuale applicazione a reti elettriche tridimensionali

3 Spectrum of the laplacian

Proposition 3.1. Variational problems related to the laplacian

Let $f \in L^2(M)$ the variational problem minimizing the Dirichlet energy functional $\int_M dx (\text{grad} f)^2$, admit as solution the kernel of the laplacian. Otherwise if we wish to have normalized functions only to avoid uniformly vanishing functions i.e. $\int_M dx f^2 = 1$ we get as solution the eigenfunction of the laplacian relative to its lowest eigenvalue. After this the same process can be applied to graphs.

Proof. The proof is splitted in two different problems:

- $\int f \Delta f = \int (\text{grad} f)^2$, the problem $\delta \int (\text{grad} f)^2 = 0$ is solved with the Euler-Lagrange equations $\text{div} \frac{\partial \mathcal{L}}{\partial (\text{grad} f)} = \frac{\partial \mathcal{L}}{\partial f}$, which give $\Delta f = 0$.
- This time we have $\delta [\int (\text{grad} f)^2 - \lambda (\int f^2 - 1)] = 0$, which using the same equations leads to $\Delta f = \lambda f$, and since for those functions $\int f \Delta f = \lambda$, we want to find the minimum eigenvalue of the laplacian.

□

Proposition 3.2. Discrete "variational" problems related to the laplacian

Let ϕ be an $k \times k$ matrix, the variational problem minimizing the Dirichlet energy $\text{Tr}(\phi^T \Delta \phi)$ is equivalent to the problem $\Delta \phi = 0$, while if we normalize the functions we get $\Delta \phi = \Lambda \phi$, where $\Lambda = \text{diag}(\lambda)$ with the k lowest eigenvalues.

Proof. As follows:

- Let $\min_\phi \text{Tr}(\phi^T \Delta \phi) = \phi_{li} \Delta_{lk} \phi_{ki} = \sum_i \phi_i^T \Delta \phi_i$ (Einstein notation) be our optimization problem, since $\Delta > 0$ we can simply solve for all ϕ_i and minimize $\frac{\partial \mathcal{L}_i}{\partial \phi_j} = \frac{\partial \phi_i^T \Delta \phi_i}{\partial \phi_j} = 2 \Delta \phi_i = 0$.
- Let $\min_\phi \text{Tr}(\phi^T \Delta \phi) = \phi_{li} \Delta_{lk} \phi_{ki} = \sum_i \phi_i^T \Delta \phi_i$ (Einstein notation) be our optimization problem, under the constraint $\phi^T \phi = I$ (orthonormal) since $\Delta > 0$ we can simply solve for all ϕ_i and minimize $\frac{\partial \mathcal{L}_i}{\partial \phi_j} = \frac{\partial (\phi_i^T \Delta \phi_i - \lambda_i (\phi_i^T \phi_i - 1))}{\partial \phi_j} = 2 \Delta \phi_i - 2 \lambda_i \phi_i = 0$. And since the initial trace is equal to the sum of our k positive eigenvalues we shall take the k lowest eigenvalues. The problem can be shown to be $\Delta \phi = \Lambda \phi$.

□