Simplicial representation learning

Learnable featurizations with neural k-forms

Kelly Maggs

10th April 2024

Collaborators

Bastian Rieck

Celia Hacker

Figure: Collaborators

Overview

Overview

Data

Set of k-simplicies $\varphi : \Delta^k \to \mathbb{R}^n$ in common ambient space

Learnable feature maps

Differential k-forms $\omega \in \Omega^k(\mathbb{R}^n)$ on the ambient space.

Featurization/representation

Integration

$$\omega : \varphi \mapsto \int_\varphi \omega \in \mathbb{R}$$

I: Motivation – Perceptrons

Perceptrons

Setup

1. **Data:** vectors $u \in \mathbb{R}^n$ i.e $u : \Delta^0 \to \mathbb{R}^n$
2. **Features:** dual vectors $v^* \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$
3. **Pairing:**

$$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}) \otimes \mathbb{R}^n \to \mathbb{R}; \qquad (v^*, u) \mapsto \langle v, u \rangle$$

Picture on board.

Perceptrons

Properties

1. *Finitely parametrized:* feature maps

$$f(v) : u \mapsto \sum v_i u_i$$

2. *Learnable:* loss gradients exist

$$\frac{\partial}{\partial v_i} \mathcal{L} f(v)(u)$$

3. *Generalizable:* defined over ambient space

$$f(v) : \mathbb{R}^n \to \mathbb{R}$$

II: Singular simplices as data

Simplicial data

Simplicial data

Dataset: Maps

$$\varphi : \Delta^n \to \mathbb{R}^n$$

into a common ambient space.

(Generalized) Simplicial data

Dataset: Singular chains

$$C_k^{sing}(\mathbb{R}^n) = \mathbb{R}\{\varphi : \Delta^k \to \mathbb{R}^n\} = \lambda_1\varphi_1 + \ldots + \lambda_k\varphi_k.$$

Question

What's a good space of feature maps for these?

Feature maps for simplices

Question

What's a good space of feature map for these?

Candidate

Singular cochains

$$C_k^{sing}(\mathbb{R}^n)^* = \mathrm{Hom}(C_k^{sing}(\mathbb{R}^n), \mathbb{R})$$

Problems

1. Vast.
2. Difficult to parametrize.
3. **Solution:** *differential forms*.

III: Differential Forms as Feature Maps

Exterior Algebra

Definition

The *free exterior algebra* $\Lambda(dx_1, \ldots, dx_n)$ is the free tensor algebra

$$T(dx_1, \ldots, dx_n)/\sim$$

with the relation $dx_i \wedge dx_j \sim -dx_j \wedge dx_i$.

$\mathbb{R}$		$\dim = 0$
dx_i	$1 \leq i \leq n$	$\dim = 1$
$dx_i \wedge dx_j$	$1 \leq i < j \leq n$	$\dim = 2$
$dx_i \wedge dx_j \wedge dx_k$	$1 \leq i < j < k \leq n$	$\dim = 3$
$\vdots$	$\vdots$	

Exterior Algebra

Intuition

Algebraic model of (infinitesimal) determinants.

$$dx_i \wedge dx_j : T_p\mathbb{R}^n \otimes T_p\mathbb{R}^n \to \mathbb{R}; \qquad \begin{matrix} \\ i \\ \\ j \\ \\ \end{matrix}\begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \mapsto \det\left(\begin{matrix} i \\ j \end{matrix}\begin{bmatrix} * & * \\ * & * \end{bmatrix} \right)$$

Note

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

Infinitesimal volume elements

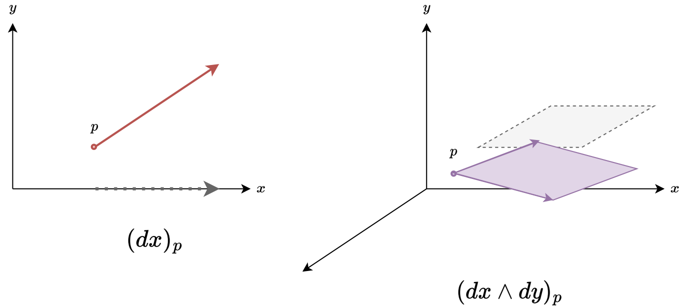


Figure: Infinitesimal determinants/volume elements.

Differential forms

Definition

The **differential forms** on $\mathbb{R}^n$ are

$$\Omega^k(\mathbb{R}^n) := C^\infty(\mathbb{R}^n) \otimes \Lambda(dx_1, \ldots, dx_n)$$

$$\begin{array}{rrl} f & & \text{0-forms} \\ f_i \cdot dx_i & 1 \leq i \leq n & \text{1-forms} \\ f_{i,j} \cdot dx_i \wedge dx_j & 1 \leq i < j \leq n & \text{2-forms} \\ f_{i,j,k} \cdot dx_i \wedge dx_j \wedge dx_k & 1 \leq i < j < k \leq n & \text{3-forms} \\ \vdots & \vdots & \end{array}$$

Differential forms

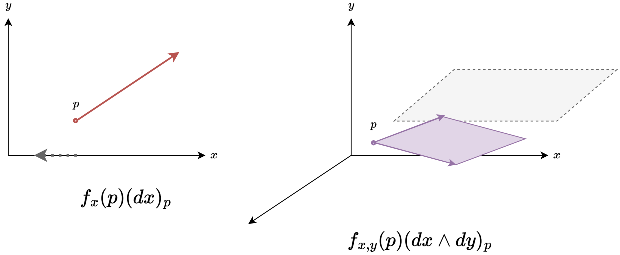


Figure: $f_I(p)$ *scales* volume in I-subspace at $p \in \mathbb{R}^n$.

Key point

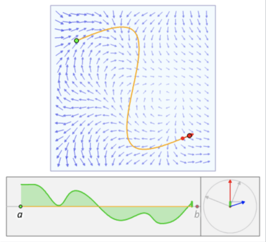
1. The k-forms $\sum_I f_I \cdot dx_I$ are determined by $f_I \ \forall I = i_1 < \ldots < i_k$
2. These functions can be parametrized/learnable

Integrating Differential Forms

Featurization

The *integral* of a k-form against a k-simplex $\varphi : \Delta^k \to \mathbb{R}^n$ is

$$\int : \Omega^k(\mathbb{R}^n) \otimes C_k^{sing}(\mathbb{R}^n) \to \mathbb{R}; \qquad (\omega, \varphi) \mapsto \int_\varphi \omega = \int_{\Delta^k} \varphi^* \omega$$



$$\longrightarrow \int_{\sigma_1} \omega \in \mathbb{R}$$

Learnable Differential form

Definition

A *neural k-form* consists of

1. $\binom{n}{k}$ MLPs
$$f_I(\mathbf{W}) : \mathbb{R}^n \to \mathbb{R}; \qquad p \mapsto \sigma\big(\mathbf{W}p\big)$$
2. A differential k-form $\omega = \sum_I f_I(\mathbf{W}) \cdot dx_I$.

Properties

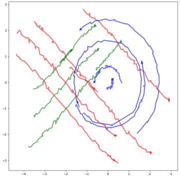
1. Finitely parametrized by **W** for each I.
2. Feature maps $\int \omega : (\varphi : \Delta^k \to \mathbb{R}^n) \mapsto \int_\varphi \omega$.
3. Gradients exist
$$\frac{\partial}{\partial \mathbf{W}} \mathcal{L}\Big(\int \omega\Big)(\varphi).$$

Experiments

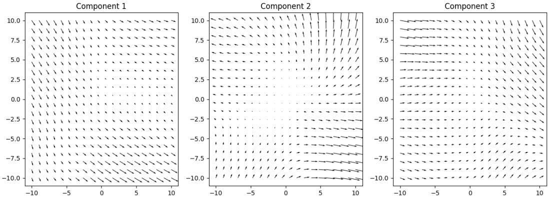
Classification of Paths

Paths in $\mathbb{R}^2$

1. Three classes of paths
2. Learn three vector fields (1-forms)
3. Integration over three vector fields $\Rightarrow$ path vectorized in $\mathbb{R}^3$.



(a) Paths.

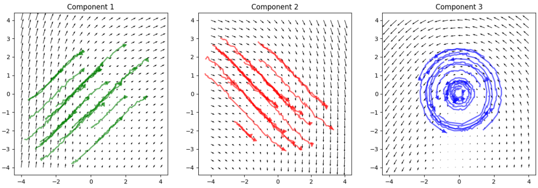


(b) Randomly initialised 1-forms.

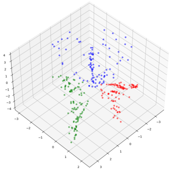
Classification of Paths

Paths in $\mathbb{R}^2$

1. Three classes of paths
2. Learn three vector fields (1-forms)
3. Integration over three vector fields $\Rightarrow$ path vectorized in $\mathbb{R}^3$.



(a) Learned 1-forms corresponding to paths in each class.



(b) Representations.

Graph Learning

Graph Learning Problems

1. Data is graphs with (non-equivariant) node embeddings.
2. Classify using learnable neural 1-forms integrated on edges.

Table: Results (mean AUROC and standard deviation of 5 runs) on benchmark datasets from the 'MoleculeNet' database Wu18a.

	Params.	BACE	BBBP	HIV
GAT	135K	69.52 ± 17.52	76.51 ± 3.36	56.38 ± 4.41
GCN	133K	66.79 ± 1.56	73.77 ± 3.30	68.70 ± 1.67
GIN	282K	42.91 ± 18.56	61.66 ± 19.47	55.28 ± 17.49
N*k*F (ours)	9K	**83.50 ± 0.55**	**86.41 ± 3.64**	**76.70 ± 2.17**