Equivarient graph convolution with Dynkin gauge quivers

Luke J Pereira

Disclaimer: This is a recreational unfinished research notebook, please contact me if you are interested in collaborating (lukejoepereira@gmail.com).

Abstract

This paper aims to extend Graph Convolution Networks (GCN) using findings from algebraic geometry and gauge theory involving Dynkin quivers and Gabriel's theorem on ADE classification. We propose the construction of a "universal" graph kernel formalized as an ADE gauge quiver, to be used for generic classification tasks with architecture and training methods inspired by diffusion on Sheaf Neural Networks (SNN). After validating this framework with experiments, we flesh out theory motivating these design choices, namely an error or residue estimation formulated in terms of a symplectic monodromy group.

Contents

1 ADE Gauge Equivariant Convolution

1.1 Overview

This paper aims to gather and organize speculative relationships between the following research:

- Neural sheaf diffusion in graph convolutional networks [**https://doi.org/10.48550/arxiv.2202.04579**]
- Gabriel's theorem on representations of ADE quivers and Gauge quivers [**nlab:quiver**],
- Nodal surplus and graph cycles via magnetic perturbation of eigenfunctions [**Berkolaiko˙2013**],
- Gromov p-widths describing non-linear spectra and isoperimetric volume bounds [**10.1007/BFb0081739**].

To begin, we claim an equivalence between Sheaf Neural Networks (§2.5) and representations of quivers (§2.1) by relaxing constraints on a sheaf connection Laplacian to no longer be strictly positive semi-definite. By allowing the connection Laplacian to have an indefinite form with positive, zero, or negative eigenvalues, we prevent the "gluing" axiom of sheaves from holding and instead have a presheaf, which can more readily be equated to a quiver representation (§2.7). In the context of classification tasks, the gluing axiom of sheaves is necessary to allow for simple diffusion by forming restriction maps between all pairwise node combinations (§2.6). However, when an indefinite connection Laplacian and disconnected presheaf takes on values from an algebraically closed field, we may instead derive an equality with quiver representations (§2.7). Then, by invoking Gabriel's Theorem, we find that *all* finite quiver representations (equivalently, all presheaf connection Laplacians) must have underlying connected subgraphs that can be classified with ADE types (§2.2). Thus, Gabriel's Theorem enables restoring a weaker gluing axiom and a modified diffusion mechanism by using tree subgraphs and the nodal surplus of cycles. In this alternative setting, the connection Laplacian becomes more akin to a graph convolution kernel (§2.8). The proposed ADE subgraphs are bipartite trees and were originally developed to describe relationships between irreducible representations (irreps) and root systems of simply connected Lie groups: SL(N+1), SU(2N), E6, E7, E8. Using spectral analysis, namely Perron-Frobenius theorem, we may classify subtrees into their ADE type using their unique bounded spectral radius $\lambda_r \leq 2$ (§2.11).

To formulate an alternative to sheaf diffusion, we need information about the non-trivial topology that results from a graphs cycles not captured by its ADE subtrees. A relevant stream of research also applies spectral methods through magnetic field perturbations to a graph Laplacian in order to compute its nodal surplus, which directly relates to a graphs cyclic structure (§2.9). Transitioning from an arbitrary graph to a collection of trees requires breaking β cycles, where $\beta = |E| - |V| + 1$ is the graph's first Betti number, which will be considered as the rank 1 perturbations. It is known that the n-th eigenfunction of a Laplacian has $(n - 1 + s)$ "zeroes", where a zero corresponds to graph edges where the eigenfunction changes sign and s is the nodal surplus or defect, which is an integer between 0 and the number of cycles. The examined method induces perturbations on a Laplacian using a magnetic field (i.e. a discretized Schrodinger operator) parameterized by its eigenvalues in what's known as a magnetic Laplacian. The fundamental result proves that the Morse index of the critical points of the perturbation field are equal to the nodal surplus of the original graph. The process of diffusion on the connection Laplacian and the application of Morse theory on critical points of the magnetic Laplacian have close similarities when studying the Hessian of each Laplacian as a mediator of covariance (§2.10).

Relating the two findings we aim to develop a graph kernel that captures a rich and invariant representation of both the graphs subtrees and its cycles. We propose a kernel that can be thought of as a histogram of the simply-connected Lie groups corresponding to ADE trees along with a symplectic monodromy group that captures the nodal surplus and their relations to the magnetic Laplacian zeroes. A monodromy group is the quotient of holonomy group (roughly, its nontrivial cycles) by the normal subgroup formed by parallel transports along homotopically trivial loops (roughly, its ADE trees). This graph kernel made of ADE trees and a symplectic monodromy group can be succinctly described as a gauge quiver or quiver diagram (§2.4). Recall that relaxing the strictly non-zero constraint on eigenvalues of the connection Laplacian leads us to Gabriel's theorem of ADE tree classification, while relaxing the constraint on linearity of an eigenvector by considering non-linear perturbations of an eigenfunction leads to topological information about a graph's nodal surplus and cyclic structure. We may examine the nonlinearity and indefiniteness of the spectra in both settings by considering a device known as a Gromov width or p-width (§2.12). This width is proposed to be interpreted as *nonlinear spectra* of a Laplacian and can be used to bound volume spectrum in an isoperimetric law. In a learning mechanism, this is used to limit the number of ADE trial graphs in a random-walk type graph kernel. This is presented in more detail in the following section.

1.2 Architecture

1.3 Experiments

2 Appendix A: Theory and Intuitions

2.1 Quiver Representations

Although it is straightforward to say that quivers are identical to directed graphs, their usefulness arises from a change in perspective that allows formulating connections between simplicial graphs and continuous topology through representation theory and categorical set theory. We may motivate this less intuitive viewpoint by recalling Grothendieck's notion of *relative point of view*, where instead of holding up individual objects, one works with families of objects or categories that depend on a creatively constructed parameterization.

Definition 2.1. Formally, a quiver is given by $Q = (Q_0, Q_1, s, t)$, where Q_0, Q_1 are finite sets with Q_0 being vertices, Q_1 being arrows corresponding to edges, and $s, t : Q_1 \to Q_0$ being maps referred to as the source and target sets of a given edge. An arrow $\alpha \in Q_1$ is written as $\alpha : s(\alpha) \to t(\alpha)$.

With this change in perspective, we allow vertices Q_0 to become fundamental while edges Q_1 become closer to categorical sets. The quiver Q can be perceived as something more akin to a point cloud than a graph; though instead of the points being embedded in a topological space like $\mathbb{R}^3$, they are embedded in an algebraic field. Edges being referred to arrows suggests a categorical parameterization, so that they relate sets (or equivalence classes) of what were formerly scalar values. An underlying graph $\bar{Q}$ can be recovered by indexing into subset of the output of adjacent pairs of surjective mappings s, t applied to an arrow $\{s(\alpha), t(\alpha)\}$.

Definition 2.2. A quiver representation $M = (M_x, M_\alpha)$ is given by vector spaces M_x for vertex $x \in Q_0$ and linear maps $M_\alpha : M_{s(\alpha)} \to M_{t(\alpha)}$.

A representation of a quiver Q is an association of an R-module to each vertex of Q and a morphism between each module for each arrow. Continuing the previous analogy, a quiver thought of as an algebraic point cloud now also takes on a topological embedding (similar to real point clouds) using its linear or fibered representations. A representation is said to be decomposable if it is isomorphic to the direct sum of non-zero representations. This notion is closely related to irreducibility of group generators and root systems,

In categorical terms, we can define a quiver to be a functor $G : X^{op} \to \mathsf{Set}$, where X^{op} is the category with objects 0, 1 and two morphisms $s, t : 1 \to 0$, along with identity morphisms. This lets us define Quiv as the category of presheaves on X, where objects are functors and morphisms are natural transformations between such functors. Then a representation of Q is a covariant functor from this category to the category of finite dimensional vector spaces. Morphisms between representations commute with arrows allowing for representations to also be considered an Abelian category.

2.2 Dynkin Diagrams

The initial ambitions of representation theory were to construct lists of all the indecomposable representations when possible, and only after to consider homomorphisms and extensions between the indecomposable objects. It turns out that the list of indecomposables is typically quite uninteresting, and instead, describing the internal categorical structure and the interplay between indecomposable representations would yield much deeper insights. In order to do so, one may look at sets of indecomposables which are related either by small changes of parameters or by the existence of irreducible maps. The interplay and binding of particular irreducible representations can be understood as an algebra or even a dynamical system in a differential setting as with Lie Groups. Then, the periodic cycles, orbits, or automorphisms are the algebraic structures that bind irreducible representations and for a group structure.

Dynkin diagrams first appeared in relation to the classification of simple Lie groups where they describe a basis of roots for a path algebra that spans a complex semi-simple Lie algebra or a compact Lie algebra and its corresponding simply laced Lie groups. P. Gabriel introduced the notion of a quiver and its representations and used them to prove the famous Gabriel's theorem on representations of quivers over algebraically closed field.

Theorem 2.1. *Let Q be a finite quiver and $\bar{Q}$ the undirected graph obtained from Q by deleting the orientation of all arrows. A connected quiver Q is of finite type if and only if the graph $\bar{Q}$ is one of the following simply laced Dynkin diagrams: A_n, D_n, E_6, E_7 or E_8.*

We can think of the diagrams as a topological group being condensed into a graph depicting interactions of its generators, which are its irreducible representations or roots. The diagram is an orthogonal sum of an irreducible root systems. Dynkin diagrams summarize relative orientations and orderings of these roots through a kaleidoscopic construction that describe its topology in terms of the group algebra.

2.3 ADE Classification

Dynkin diagrams have the following correspondence with the Lie algebras associated to classical groups over the complex numbers, ADE types have additional compact Lie algebras and corresponding simply laced Lie groups:

- A_n: $\mathfrak{sl}_{n+1}(\mathbb{C})$, the special linear Lie algebra of traceless operators. Also corresponds to $\mathfrak{su}_{n+1}(\mathbb{R})$, the algebra of the special unitary group $SU(n+1)$.
- B_n: $\mathfrak{so}_{2n+1}(\mathbb{C})$, the odd-dimensional special orthogonal Lie algebra.
- C_n: $\mathfrak{sp}_{2n}(\mathbb{C})$, the symplectic Lie algebra.
- D_n: $\mathfrak{so}_{2n}(\mathbb{C})$, the even-dimensional special orthogonal Lie algebra ($n > 1$) of even-dimensional skew-symmetric operators. Also corresponds to $\mathfrak{so}_{2n}(\mathbb{R})$, the algebra of the even projective special orthogonal group $PSO(2n)$
- E_6, E_7, E_8: the names for the exceptional Lie groups and algebras coincide with the associated Dynkin diagram.

The graphs describes a finite reflection group with each node representing a reflection satisfying relations depicted as (labeled) edges. The edges in the graphs show that two fundamental roots are not orthogonal (perpendicular) but differ by 120 degrees or $2\pi/3$. We can consider repeated reflective action as an exponential rotation of $(2\pi/3)^k$ that yields equivalences between self or pairs. These self or pairwise interactions with exponents of 2 have no edge. Interactions with exponents of 3 have labels omitted. Repeated reflections resulting in the identity (periodic automorphisms) are shown to be equivalent to commutativity between pairs of generators. Conjugation invariance (like the reflection periodicity) is also equivalent to commutation. The normal subgroup, which is an equivalance class of the identity, is also the center of an orbit and can also be understood as measure of commutativity. Inner automorphisms measure failure/divergence from commutativity, outer automorphisms measures non-inner automorphisms and are isomorphic to automorphisms of Dynkin diagrams.

2.4 Gauge Quivers

Definition 2.3. A quiver gauge theory is given by the following:

- Finite quiver Q
- Each vertex $v \in \mathrm{V}(Q)$ corresponds to a compact Lie group G_v. This may be the unitary group $U(N)$, the special unitary group $SU(N)$, special orthogonal group $SO(N)$ or symplectic group $USp(N)$ corresponding to ADE classes.
- The gauge group is the product $\prod_{v\in \mathrm{V}(Q)} G_v$.
- Each edge of Q, $e\colon u \to v$, corresponds to the defining representation $\bar{N}_u \otimes N_v$. This representation is called a bifundamental representation.

The quiver is particularly convenient for representing conformal gauge theory.

2.5 Sheaf Neural Networks

A Sheaf Neural Network is a type of Graph Neural Network that operates on a sheaf, an object that equips a graph with vector spaces over its nodes and edges and linear maps between these spaces.

Definition 2.4. A cellular sheaf $(G, \mathcal{F})$ on an undirected graph $G = (V, E)$ consists of:

- A vector space $\mathcal{F}(v)$ for each $v \in V$,
- A vector space $\mathcal{F}(e)$ for each $e \in E$,
- A linear map $\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \to \mathcal{F}(e)$ for each incident node-edge pair $v \trianglelefteq e$.

This definition closely resembles that of the quiver representation, though there is an additional vector space equipped to each edge and linear maps explicitly given between each pair of nodes. This allows the presheaf corresponding to the quiver representation to be promoted into a sheaf by satisfying an additional "gluing" axiom.

The vector spaces of the node and edges are called stalks, while the linear maps are called restriction maps. It is possible to group the various spaces by interpreting the graph as a 1-dimensional simplicial complex. In this setting, the 0-dimensional simplicies correspond to nodes and the 1-dimensional simplicies are edges. A p-chain of a simplicial complex is the sum of its p dimensional simplicies. For a graph, the 0-chains are aggregation of nodes and 1-chains are aggregations of edges. Likewise, the dual space formed by the node stalks is called the space of 0-cochains, while the dual space formed by edge stalks is called the space of 1-cochains.

Definition 2.5. Given a sheaf $(G, \mathcal{F})$, we define the space of 0-cochains $C^0(G, F)$ as the direct sum over the vertex stalks $C^0(G, F) := \oplus_{v \in V} \mathcal{F}(v)$. Similarly, the space of 1-cochains $C^1(G, F)$ as the direct sum over the edge stalks $C^1(G, F) := \oplus_{e \in E} \mathcal{F}(e)$.

Definition 2.6. Given some arbitrary orientation for each edge $e = u \to v, e \in E$, we define the coboundary map $\delta : C^0(G, F) \to C^1(G, F)$ as $\delta(x)_e = \mathcal{F}_{v \trianglelefteq e} x_v - \mathcal{F}_{u \trianglelefteq e} x_u$. Here $x \in C^0(G, F)$ is a 0-cochain and $x_v \in \mathcal{F}(v)$ is the vector of x at the node stalk $\mathcal{F}(v)$.

A p-boundary is considered to be the aggregation of p-simplicies in a p-chain that takes p-chains to $p+1$-chains. A p-coboundary is a dual homomorphism that that takes p-cochains to $p+1$-cochains. From an opinion dynamics perspective (Hansen & Ghrist, 2021), the node stalks may be thought of as the private space of opinions and the edge stalks as the space in which these opinions are shared in a public discourse space. The coboundary map δ then measures the disagreement between all the nodes.

A p-cycle describes a loop resulting from a closed p-chain. A homology group is defined by quotienting the group of p-cycles Z_p by the group of p-boundaries B_p, i.e. $H_p = Z_p/B_p$. Similarly the cohomology group can be defined as a quotient of p-cocycles and p-coboundaries, i.e. $H^p = Z^p/B^p$. This will be relevant in later sections involving symplectic cyclic geometry. Recall the previous notion of Dynkin quiver representations having $Ext^2 = 0$, this would be equivalent to having empty cohomology, meaning the group of coboundaries would be infinite or the group of cycles is empty. Conversely, we may use this restriction to categorize the manner in which the data diverges from being acyclic or a tree quiver.

The sheaf Laplacian operator is a symmetric positive semi-definite block matrix resulting from multiplication of the the co-boundary operator by its transpose. By making the sheaf laplacian symmetric, we force the "gluing" axiom to be true, and enable the pre-sheaf to become a sheaf. This also means the sheaf has a similar undirected nature as the underlying graph and makes spectral analysis possible.

Definition 2.7. The sheaf Laplacian of a sheaf is a map $\mathcal{L}_{\mathcal{F}} : C^0(G, \mathcal{F}) \to C^0(G, \mathcal{F})$ defined as $\mathcal{L}_{\mathcal{F}} = \delta^T \delta$. The normalised sheaf Laplacian $\Delta\mathcal{F}$ is defined as $\Delta\mathcal{F} = D^{-\frac{1}{2}} \mathcal{L}_{\mathcal{F}} D^{-\frac{1}{2}}$ where D is the blockdiagonal of $\mathcal{L}_F$.

If we constrain the restriction maps in the sheaf to belong to the orthogonal group, the sheaf becomes a discrete $O(d)$-bundle and can be thought of as a discretised version of a tangent bundle on a manifold. The sheaf Laplacian of the $O(d)$-bundle is equivalent to a connection Laplacian used by Singer & Wu (2012). The orthogonal restriction maps describe how vectors are rotated when transported between stalks, in a way analogous to the transportation of tangent vectors on a manifold.

2.6 Neural Sheaf Diffusion

Consider a graph $G = (V, E)$ where each node $v \in V$ has a d-dimensional feature vector $x_v \in \mathcal{F}(v)$. We construct an nd-dimensional vector $x \in C^0(G, F)$ by column-stacking the individual vectors x_v. Allowing for f feature channels, we produce the feature matrix $X \in \mathbb{R}^{(nd)\times f}$. The columns of X are vectors in $C^0(G, F)$, one for each of the f channels. Sheaf diffusion is a process on (G, F) governed by the following discreted diffusion equation:

$$X_{t+1} = X_t - \sigma(\Delta_{\mathcal{F}(t)} I_n \otimes W_1^t) X_t W_2^t \tag{1}$$

It is important to note that the sheaf $\mathcal{F}(t)$ and the weights W_1^t, W_2^t are time-dependent, meaning that the underlying "geometry" evolves over time. The diffusion of features into the kernel of the laplacian can be understood as convolution of adjacent nodes in which the sheaf serves as a multi-headed attention mechanism.

By relaxing constraints on the sheaf Laplacian being symmetric positive semi-definite and omitting constraints on the restriction maps, we can construct an equivalence between a pre-sheaf neural network and a quiver representation. To maintain the symmetry of the sheaf we only require that that its underlying graph be strongly connected.

2.7 Equivalence of presheaves and quiver representations

2.8 Graph convolution kernel

2.9 Magnetic Laplacian and nodal surplus

2.10 Correspondences from Hessian

2.11 Correspondences from spectral analysis

Theorem 2.2. *Perron-Frobenius theorem tells us that if our graph or subgraph is strongly connected, then its Laplacian (which must be a non-negative irreducible matrix) will have the form ωr where r is a real strictly positive eigenvalue, and ω ranges over the complex h-th roots of unity for some positive integer h called the period of the matrix.*

Theorem 2.3. *Let G be a finite simple graph (without loops or multiple edges) and denote its spectral radius r_G. Then $r_G < 2$ if and only if each connected component of G is one of Dynkin diagrams A_n, D_n, E_6, E_7, E_8. Moreover, $r_G = 2$ if and only if each connected component of G is one of the extended Dynkin diagram $\tilde{A}_n, \tilde{D}_n, \tilde{E}_6, \tilde{E}_7, \tilde{E}_8$*

2.12 Gromov width and isoperimetric bounds

3 Appendix B: Future Research

3.1 Automorphisms

As seen in Gabriel's theorem, any finite representation must have underlying ADE graphs. Automorphisms of ADE diagrams are equivalent to the outer Automorphism group which composes with the inner automorphism group that represents a measure of noncommutativity (non-abelian) of the group. Dynkin quivers can't have any indecomposable quiver representations nor have automorphisms other than scalars, nor any self-extensions. Neural sheafs may be more akin to derived category of coherent sheaves (on a smooth algebraic or projective variety and and on their noncommutative counterparts). Recall, the cohomology group can be defined as a quotient of p-cocycles and p-coboundaries, i.e. $H^p = Z^p/B^p$. As seen in sheaf neural networks, this can be related to the sheaf Laplacian constructed from coboundaries of cellular or simplicial complexes.

3.2 Extensions, Filtrations, and Cohomology

The category of quiver representations over a field is hereditary, with $\mathsf{Ext}^2(M, N) = 0$ for any representations M, N. The extensions Ext are the derived homs, meaning they are homs not of modules but of chain complexes, and are exact in that they preserve quasi-isomorphism.

Definition 3.1. Let H be a finite-dimensional algebra and $S(1), \ldots, S(N)$ be the simple modules of H corresponding to irreducible representations. Let Q be the Ext-quiver of H, i.e. Q has as vertices the simple modules $S(1), \ldots, S(N)$ and an arrow $S(i) \to S(j)$ provided $\mathsf{Ext}^1_H(S(i), S(j)) \neq 0$.

in finite global dimension there cannot be a loop in the Ext-quiver

Relates to Ext functor being 0