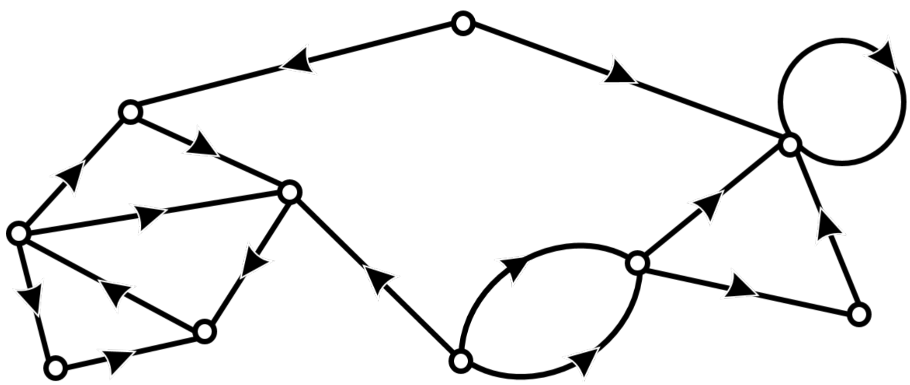
Principal Components along Quiver Representations

Anna Seigal, Heather A. Harrington, and Vidit Nanda

ABSTRACT. Quiver representations arise naturally in many areas across mathematics. Here we describe an algorithm for calculating the vector space of sections, or compatible assignments of vectors to vertices, of any finite-dimensional representation of a finite quiver. Consequently, we are able to define and compute principal components with respect to quiver representations. These principal components are solutions to constrained optimisation problems defined over the space of sections, and are eigenvectors of an associated matrix pencil.

Introduction

A **quiver representation** is an arrangement of vector spaces and linear maps tethered to the vertices and edges of a directed graph [**DW17**, **Sch14**]. The quiver illustrated below will be our running example throughout the paper.



Despite being relatively concrete mathematical objects, quiver representations provide a uniform framework for a host of fundamental abstract problems in linear algebra [**BGP73**]. They also play an important role in various other fields, including the study of associative algebras [**ARS97**], Gromov-Witten invariants [**GP10**], representations of Kac-Moody algebras [**Nak01**], moduli stacks [**Tod18**], Morse theory [**KP19**], persistent homology [**Oud15**] and perverse sheaves [**GMV96**], among others.

In most of these contexts, the crucial property of a given quiver representation is its *decomposability* into a direct sum of smaller representations. Gabriel's celebrated result [**Gab72**] establishes that a quiver admits finitely many (isomorphism classes of) indecomposable representations if and only if its underlying undirected graph is a simply-laced Dynkin diagram (i.e., type A, D or E). Thus, most quivers have rather complicated sets of indecomposable representations, and are said to be of *wild* type. It is a direct consequence of this trifecta – concreteness, ubiquity and generic wildness – that ideas from disparate branches of mathematics have conversely been deployed to study representations of quivers. These include algebraic geometry [**Kir16**], combinatorics [**DW11**], differential geometry [**HW11, HKKP17**], geometric representation theory [**Gin09**], invariant theory [**Kin94**, **Kac80**], and multilinear algebra [**Her08**].

Quiver representations have recently emerged in far more applied and computational contexts than the classical ones listed above. We are aware of four such appearances: the first is in the study of *cellular sheaves* [**Cur13**], which are representations of (the Hasse graph associated to) a finite regular cell complex. The second appearance takes the form of *connection matrices* in computational dynamics — these are indexed by the Conley-Morse graph of a discretised dynamical system, with vertices corresponding to recurrent sets and edges representing gradient flow [**HMS18**, Def 9.19]. The third appearance is in the algebro-geometric study of neural network architectures [**AJ20**, **JL21**]. And finally, (stability properties of) quiver representations play a role in algebraic statistics [**AKRS20**, **DM20**]. We expect (and hope) that this influx of quiver theory into more applied domains will continue.

This Paper. We consider a representation $\mathbf{A}_\bullet$ of a quiver Q. It assigns vector spaces $\mathbf{A}_v$ to each vertex v of Q and linear maps $\mathbf{A}_e$ to each edge e in Q. We construct a vector space $\Gamma(Q;\mathbf{A}_\bullet)$ called the **space of sections** of the quiver representation. An element of $\Gamma(Q;\mathbf{A}_\bullet)$ selects one vector γ_v from the vector space $\mathbf{A}_v$ assigned to each vertex v so that for every edge $e : u \to v$ the linear map $\mathbf{A}_e$ sends γ_u to γ_v. As such, $\Gamma(Q;\mathbf{A}_\bullet)$ is a subspace of the *total space* $\mathrm{Tot}(\mathbf{A}_\bullet) := \prod_v \mathbf{A}_v$. The assignment

$$\mathbf{A}_\bullet \mapsto \Gamma(Q;\mathbf{A}_\bullet)$$

can directly be seen to be a functor from the category of Q-representations to the category of vector spaces. We do not expect this functor to directly answer any deep questions regarding (in)decomposability of quiver representations. Rather, we hope that the space of sections will become a useful and practical tool for those who encounter quiver representations in applied and computational contexts.

Our first contribution is an algorithm for computing the space of sections for any finite-dimensional representation of a finite quiver. This is of some relevance even to those who have no warm feelings for quiver representations, since it is a purely categorical procedure for computing the limit (i.e., the universal cone) of a diagram in the category of vector spaces. With minor modifications, it can be made to work for diagrams valued in any abelian category that has computable products and equalisers. There are two major difficulties associated with calculating the space of sections: the first of these arises from directed cycles, where we are forced to restrict to a fixed point space of an endomorphism; and the second is the presence of multiple incoming edges at a vertex, where we are forced to restrict to an equaliser.

None of these difficulties arise when the quiver is a directed rooted tree. Thus, our algorithm consists of two steps — the first step removes all directed cycles and updates the representation $\mathbf{A}_\bullet$ accordingly; and the second step replaces this acyclic quiver with a directed rooted tree, again updating the representation. The result is a new representation $\mathbf{A}^+_\bullet$ of a rooted directed tree T^+, which has all the same vertices as Q (plus an additional root vertex) and satisfies $\mathbf{A}^+_v \subset \mathbf{A}_v$ at each vertex v.

Here is our first main result.

THEOREM (A). *The space of sections* $\Gamma(Q;\mathbf{A}_\bullet)$ *is the image of the map*

$$F : \mathbf{A}^+_r \longrightarrow \mathrm{Tot}(\mathbf{A}_\bullet),$$

obtained by composing the linear maps assigned by the quiver representation $\mathbf{A}_\bullet^+$ *along the unique path in the rooted directed tree* T^+ *from the root* r *to each other vertex.*

Although the constructions of T^+ and $\mathbf{A}_\bullet^+$ are explicit and readily implementable on a computer, they require making several intermediate choices. Each such choice is liable to produce a different F, but its image is always $\Gamma(Q; \mathbf{A}_\bullet)$ regardless of these choices.

Our second contribution takes place in the realm of quiver representations valued in real vector spaces; in this case, a map F as described in Theorem (A) can be represented by an $n \times d$ full-rank real matrix, where n and d are the dimensions of $\mathrm{Tot}(\mathbf{A}_\bullet)$ and $\Gamma(Q; \mathbf{A}_\bullet)$ respectively. Using this matrix, we define the **principal components** of any (generic, mean-centered) finite set D of vectors in $\mathbb{R}^n \simeq \mathrm{Tot}(\mathbf{A}_\bullet)$ with respect to the quiver representation $\mathbf{A}_\bullet$. As with ordinary principal components, the starting point is the $n \times n$ sample covariance matrix S of the vectors in D. Next, we consider for each $r \leq d$ the variational problem of maximising the trace $\mathrm{tr}(X^\mathsf{T} S X)$ over the set of all $n \times r$ matrices X that satisfy $X^\mathsf{T} X = \mathrm{id}$, and whose columns are constrained to lie in the image of F, i.e., in $\Gamma(Q; \mathbf{A}_\bullet)$. There is a unique solution, obtained by iteratively incrementing r from 1 to d, and the span of its r-th column is the r-th principal component of D along $\mathbf{A}_\bullet$, denoted $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$. Unlike ordinary principal components, the $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ are not spanned by eigenvectors of S in general.

The second main result of this paper is that the quiver principal components $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ do in fact admit a spectral interpretation.

THEOREM (B). *For each* $1 \leq r \leq d$, *the r-th principal component* $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ *is spanned by* Fu_r, *where* u_r *is the eigenvector of the matrix pencil* $F^\mathsf{T} S F - \lambda(F^\mathsf{T} F)$ *corresponding to its r-th largest eigenvalue.*

We see from the matrix pencil in Theorem (B) that the principal components along a quiver representation intertwine the properties of D (via the sample covariance matris S) with those of quiver Q (via the map F to the space of sections). These principal components find directions of maximum variation among vectors in $D \subset \mathbb{R}^n$ that respect certain linear dependencies. The coordinates in $\mathbb{R}^n$ can be thought of as partitioned into blocks (one per vertex of the quiver); these blocks are related by the linear maps of the quiver representation. In this way, the principal components along the quiver representation interpolate between concatenating ordinary principal components from individual blocks, and ordinary principal components in the whole space $\mathbb{R}^n$. We will mostly assume that the linear maps are fixed in advance, but we will briefly discuss approaches to learning them from the set D.

Related Work. The first half of this work is inspired by the study of cellular sheaves, which functorially assign vector spaces to cells and linear maps to incidence relations in a finite cell complex. The space of sections of a cellular sheaf $\mathscr{S}$ defined over an undirected graph G is isomorphic to the zeroth sheaf cohomology group $\mathbf{H}^0(G; \mathscr{S})$, which is readily computable [CGN16]. We can turn any representation $\mathbf{A}_\bullet$ of a quiver Q into a cellular sheaf over the underlying undirected graph by replacing each edge-indexed linear map

$$\mathbf{A}_u \xrightarrow{\mathbf{A}_e} \mathbf{A}_v$$

by a corresponding zigzag of the form

$$\mathbf{A}_u \xrightarrow{\mathbf{A}_e} \mathbf{A}_v \xleftarrow{\mathrm{id}} \mathbf{A}_v.$$

Thus, each edge inherits the vector space assigned to its target vertex. Computing zeroth cohomology of this sheaf furnishes an alternative to Theorem (A) for calculating $\Gamma(Q; \mathbf{A}_\bullet)$. However, this cohomological alternative suffers from two significant drawbacks — first, the insertion of these zigzags is quite inconvenient for our purposes of testing compatibility of sections across directed paths in the original quiver. And second, the duplication of vector spaces over the edges leads to unnecessarily large matrices, and hence incurs a larger computational cost.

A central focus of the second half of this paper is the study of linearly constrained principal components, which dates back at least to [**Rao64**, Section 11]. It is referred to as *constrained PCA* in [**DK96**, Section 7.1] and [**TH01**, **TS91**]. Its statistical implications are discussed in [**HT02**] and [**TH01**, Section 5.4]. For an example of constrained PCA occurring in a biological context, see [**Hua20**]. It is also important to note that the principal components $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ introduced in this paper do not constitute a low-rank approximation of the representation $\mathbf{A}_\bullet$. Such approximation of related multi-linear objects appears in [**BHS$^+$19**], where the authors find the singular value decomposition of a finite chain complex, and [**HMR20**], in the study of orthogonal decomposition of tensor networks. A study of star quivers for parameter estimation in integrated PCA [**TA18**] appears in [**FM21**].

Organisation. The remainder of this paper is divided into eight short sections. In §1, we define quiver representations, their sections, and some elementary properties thereof.

§2 is devoted to the task of using the ear decomposition to compute the sections of strongly-connected quivers. §3 uses the results of §2 to construct, from any given quiver representation, a sub-representation of an acyclic subquiver that has the same space of sections. In §4 we describe how to further modify this acyclic subquiver into a rooted directed tree and update the overlaid representation to preserve the space of sections. These intermediate results are assembled in §5 to provide a proof of Theorem (A); we also give lower bounds on the dimension of the space of sections.

Principal components along quiver representations are defined in §6 via three optimisation problems; we show here that all three give the same answer. In §7 we use a generalisation of the singular value decomposition to establish Theorem (B). And finally, §8 discusses the problem of learning the linear maps of a quiver representation from finite samples of vectors living in the total space.

Acknowledgements. AS and HAH thank Julian Knight, and AS thanks Visu Makam, for helpful discussions. VN is grateful to Frances Kirwan for timely quiver-theoretic advice. AS, HAH, and VN are members of the UK Centre for Topological Data Analysis and thank the Mathematical Institute at the University of Oxford. HAH and VN acknowledge funding from EPSRC grant EP/R018472/1. HAH gratefully acknowledges funding from EPSRC EP/R005125/1 and EP/T001968/1, the Royal Society RGF\EA\201074 and UF150238, and Emerson Collective.

1. Quiver Representations and Sections

A quiver Q consists of a finite set V whose elements are called vertices, a finite set E whose elements are called edges, and two maps $s, t : E \to V$ called the source and target map respectively. It is customary to illustrate quivers by drawing points for vertices and arrows (from source to target) for edges. A path in Q is an ordered finite sequence of distinct edges $p = (e_1, e_2, \ldots, e_k)$ with disjoint sources (i.e., $s(e_i) \neq s(e_j)$ when $i \neq j$) so that $s(e_{i+1}) = t(e_i)$

holds for every $1 \leq i < k$:

$$\bullet \xrightarrow{e_1} \bullet \xrightarrow{e_2} \cdots \xrightarrow{e_{k-1}} \bullet \xrightarrow{e_k} \bullet$$

The source and target maps extend from edges to paths via $s(p) = s(e_1)$ and $t(p) = t(e_k)$. We call p a *cycle* if $s(p) = t(p)$, and call Q *acyclic* if it does not admit any cycles.

A **representation** of Q comprises an assignment $\mathbf{A}_\bullet$ of a finite-dimensional vector space $\mathbf{A}_v$ to every vertex v in V and a linear map $\mathbf{A}_e : \mathbf{A}_{s(e)} \to \mathbf{A}_{t(e)}$ to every edge e in E. We will remain agnostic to the choice of underlying field until Section 6. Using the data of $\mathbf{A}_\bullet$, one can associate to each path $p = (e_1, \ldots, e_k)$ the map $\mathbf{A}_p : \mathbf{A}_{s(p)} \to \mathbf{A}_{t(p)}$ via

$$\mathbf{A}_p := \mathbf{A}_{e_k} \circ \mathbf{A}_{e_{k-1}} \circ \cdots \circ \mathbf{A}_{e_2} \circ \mathbf{A}_{e_1}. \tag{1}$$

The *total space* of $\mathbf{A}_\bullet$ is the direct product

$$\mathrm{Tot}(\mathbf{A}_\bullet) := \prod_{v \in V} \mathbf{A}_v.$$

The following terminology has been borrowed from analogous notions that arise in the study of sheaves and vector bundles.

DEFINITION 1.1. Let $\mathbf{A}_\bullet$ be a representation of a quiver $Q = (s, t : E \to V)$. A **section** of $\mathbf{A}_\bullet$ is a an element $\gamma = \{\gamma_v \in \mathbf{A}_v \mid v \in V\}$ in $\mathrm{Tot}(\mathbf{A}_\bullet)$ satisfying the *compatibility* requirement $\gamma_{t(e)} = \mathbf{A}_e(\gamma_{s(e)})$ across each edge e in E.

The set of all sections of $\mathbf{A}_\bullet$ is a vector subspace of $\mathrm{Tot}(\mathbf{A}_\bullet)$, which we denote by $\Gamma(Q; \mathbf{A}_\bullet)$. The explicit computation of the space of sections $\Gamma(Q; \mathbf{A}_\bullet)$, for any quiver Q and representation $\mathbf{A}_\bullet$, is one of the central objectives of this work.

REMARK 1.2. The product of general linear groups $\mathrm{G} = \prod_v \mathrm{GL}(\mathbf{A}_v)$ acts on $\mathbf{A}_\bullet$ by change of basis: given any $g = \{g_v \in \mathbf{A}_v \mid v \in V\}$, the new representation $g\mathbf{A}_\bullet$ assigns

$$(g\mathbf{A})_v = \mathbf{A}_v \qquad \text{and} \qquad (g\mathbf{A})_e = g_{t(e)} \circ \mathbf{A}_e \circ g_{s(e)}^{-1}.$$

This action descends to the space of sections via $\gamma \mapsto g\gamma$, where $(g\gamma)_v = g_v\gamma_v$, and so we have an isomorphism $\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(Q; g\mathbf{A}_\bullet)$ for every $g \in \mathrm{G}$. In fact, a purely formal argument shows that the assignment $\mathbf{A}_\bullet \mapsto \Gamma(Q; \mathbf{A}_\bullet)$ is a functor from the category of representations of a fixed Q to the category of vector spaces. Here a morphism $\mathscr{F}_\bullet : \mathbf{A}_\bullet \to \mathbf{A}'_\bullet$ of Q-representations is a collection of V-indexed linear maps $\mathscr{F}_v : \mathbf{A}_v \to \mathbf{A}'_v$ which commute with the edge-maps, i.e., for each $e \in E$ we have

$$\mathbf{A}'_e \circ \mathscr{F}_{s(e)} = \mathscr{F}_{t(e)} \circ \mathbf{A}_e. \tag{2}$$

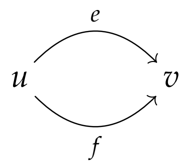
Each section $\gamma \in \Gamma(Q; \mathbf{A})$ is sent by $\mathscr{F}_\bullet$ to a section $\mathscr{F}\gamma$ of $\mathbf{A}'_\bullet$ prescribed by $(\mathscr{F}\gamma)_v = \mathscr{F}_v(\gamma_v)$, since applying (2) to $\gamma_{s(e)}$ gives the desired compatibility across each edge e:

$$\mathbf{A}'_e \circ \mathscr{F}_{s(e)}(\gamma_{s(e)}) = \mathscr{F}_{t(e)} \circ \mathbf{A}_e(\gamma_{s(e)}) = \mathscr{F}_{t(e)}\gamma_{t(e)}.$$

Compatibility across edges imposes severe constraints on sections, even in the simplest of examples, when the underlying quiver Q contains cycles or vertices with multiple incoming edges.

EXAMPLE 1.3. Consider the quiver that consists of a single vertex v and a single edge e with $s(e) = v = t(e)$. The space of sections of any representation $\mathbf{A}_\bullet$ is the subspace of $\mathbf{A}_v$ fixed by $\mathbf{A}_e$, i.e., the eigenspace corresponding to eigenvalue 1.

EXAMPLE 1.4. The space of sections of a representation $\mathbf{A}_\bullet$ of the 2-Kronecker quiver, pictured below, is isomorphic to $\ker(\mathbf{A}_e - \mathbf{A}_f)$.



In sharp contrast, sections are far less constrained when the vertices of Q admit at most one incoming edge.

EXAMPLE 1.5. Given vector spaces U, V, W along with linear maps $A : V \to U$ and $B : V \to W$, the sections of the quiver representation

$$U \xleftarrow{\;A\;} V \xrightarrow{\;B\;} W.$$

are triples of the form $\gamma = (Ax, x, Bx)$ for x in V.

More generally, consider the case where Q admits a distinguished vertex r in V called the *root* so that for each other vertex $v \neq r$ there is a unique path $p[v]$ in Q from r to v. Quivers satisfying this unique path property are studied in various contexts and hence have many names — these include *out-trees*, *out-branchings*, *directed rooted trees* and (the far more scenic) **arborescences** [BJG09, Chapter 9].

PROPOSITION 1.6. *Let $\mathbf{A}_\bullet$ be a representation of an arborescence Q with root vertex r. The space of sections $\Gamma(Q; \mathbf{A}_\bullet)$ is isomorphic to $\mathbf{A}_r$, with every section γ uniquely determined by the vector $x = \gamma_r$ in $\mathbf{A}_r$, via*

$$\gamma_v = \mathbf{A}_{p[v]}(x),$$

where $p[v]$ is the unique path in Q from r to $v \neq r$.

2. Sections of Strongly Connected Quivers

A quiver $Q = (s, t : E \to V)$ is called **strongly connected** if for any pair of vertices v, v' in V there is at least one path from v to v'. The simplest examples of strongly connected quivers are cycles, but such quivers can be far more intricate. In this section, we study sections of strongly connected quivers. We will use a particular decomposition of such quivers into a union of simpler quivers. To this end, note that a **subquiver** $Q' \subset Q$ is a choice of subsets $V' \subset V$ and $E' \subset E$ so that the restrictions of s and t to E' take values in V'. For example, every path $(e_1, \ldots, e_k)$ in Q forms a subquiver with

$$V' = \{s(e_1),\, s(e_2),\, \ldots,\, s(e_k),\, t(e_k)\} \text{ and } E' = \{e_1, \ldots, e_k\},$$

where $[k] = \{1, \ldots, k\}$. In the special case where a subquiver Q' comes from a path in Q, we define its source and tail $s(Q')$ and $t(Q')$ to be the source and tail of that path. Here is the decomposition of interest [BJG09, Sec 5.3].

DEFINITION 2.1. An **ear decomposition** $Q_\bullet$ of Q is an ordered sequence of $c \geq 1$ subquivers $\{Q_i = (s_i, t_i : E_i \to V_i) \mid i \in [c]\}$ of Q subject to the following axioms:

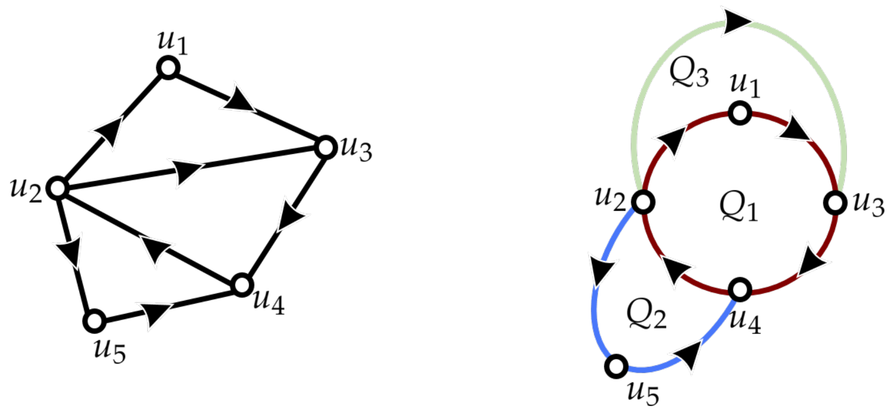
(1) the edge sets E_i partition E — in other words, they are mutually disjoint and their union equals E; moreover,
(2) the quiver Q_1 is either a single vertex or a cycle, while Q_i for each $i > 1$ is a (possibly cyclic) path in Q; and finally,

(3) for each $i > 1$, the intersection of V_i with the union $\bigcup_{j<i} V_j$ equals $\{s(Q_i), t(Q_i)\}$; this intersection has cardinality 1 if Q_i is a cycle and cardinality 2 otherwise.

Ear decompositions play an important role in the study of strongly connected quivers due to the following fundamental result.

THEOREM 2.2. *A quiver with at least two vertices is strongly connected if and only if it has an ear decomposition.*

At least one standard proof of this result is given in the form of an efficient algorithm for constructing ear decompositions — see [**BJG09**, Theorem 5.3.2] for details. The figure below illustrates a strongly connected subquiver of the quiver depicted in the Introduction along with its decomposition into three ears:



We assume for the remainder of this section that Q is strongly connected and fix an ear decomposition $Q_\bullet$ as in Definition 2.1. The *depth* of an edge $e \in E$, denoted $|e|$, is the unique $i \in [c]$ with $e \in E_i$. We say that a path $p = (e_1, \ldots, e_k)$ is $Q_\bullet$-increasing if the $|e_i|$ form a weakly increasing sequence, and define $Q_\bullet$-decreasing paths analogously. For each vertex $v \in V$, we write $\ell(v)$ for the smallest i in $[c]$ such that $v \in V_i$.

PROPOSITION 2.3. *Let r be any vertex in V_1. For any vertex $v \neq r$ in V, there exists*

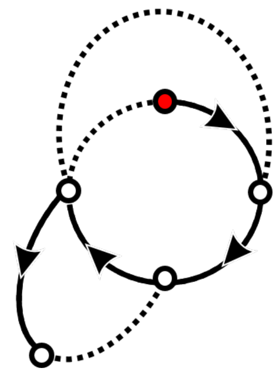
(1) *a unique $Q_\bullet$-increasing path $p[v]$ from r to v with all edges of depth $\leq \ell(v)$, and.*
(2) *a unique $Q_\bullet$-decreasing path $q[v]$ from v to r with all edges of depth $\leq \ell(v)$.*

PROOF. For $\ell = 1$, the desired conclusion follows immediately because Q_1 must be a cycle by axiom (2) of Definition 2.1. Proceeding inductively, we assume that the assertion holds whenever $\ell < i$, and consider any $v \in V_i$. Once again by axiom (2), our vertex v lies on a path Q_i from $s(Q_i)$ to $t(Q_i)$; and by axiom (3), the inductive hypothesis applies to both $s(Q_i)$ and $t(Q_i)$. The increasing path $p[v]$ is built by first going from r to $s(Q_i)$ along $p[s(Q_i)]$ and then onward to v along Q_i. Similarly, the decreasing path $q[v]$ is built by concatenating the piece of Q_i which goes from v to $t(Q_i)$ with the path $q[t(Q_i)]$. □

For each i in $[c]$, the set E_i contains at most one edge $\epsilon_i \in E_i$ whose target is $t(Q_i)$; we allow for the possibility that ϵ_1 does not exist if Q_1 has no edges, but all other ϵ_i exist and are uniquely determined by the ear decomposition. We call ϵ_i the *i-th terminal edge* with respect to the ear decomposition $Q_\bullet$, and denote the set of all terminal edges by $E_{\text{ter}} \subset E$.

DEFINITION 2.4. The **arborescence induced by** $Q_\bullet$ is the subquiver $T = T(Q_\bullet)$ with vertex set V and edges $E - E_{\text{ter}}$.

To confirm that T is an arborescence, note that its root vertex is $r = s(Q_1)$, and that for any other vertex v there is a unique path $p[v]$ from r to v, whose existence is guaranteed by Proposition 2.3. In the ear decomposition drawn above, the three terminal edges (with respect to the root vertex u_1) are replaced by dotted arcs in the figure below. The arborescence induced by $Q_\bullet$ is obtained by removing these three edges:



Given a terminal edge $\epsilon \in E_{\text{ter}}$, consider the linear map $\Delta_\epsilon : \mathbf{A}_r \to \mathbf{A}_{t(\epsilon)}$ given by

$$\Delta_\epsilon = \mathbf{A}_{p[t(\epsilon)]} - \mathbf{A}_\epsilon \circ \mathbf{A}_{p[s(\epsilon)]}. \tag{3}$$

The kernel of each such map is a subspace $\ker \Delta_\epsilon \subset \mathbf{A}_r$. These kernels depend on the choice of ear decomposition $Q_\bullet$ and the representation $\mathbf{A}_\bullet$. Let's denote their intersection by

$$K(Q_\bullet; \mathbf{A}_\bullet) := \bigcap_\epsilon \ker \Delta_\epsilon, \tag{4}$$

where ϵ ranges over E_{ter}. Up to isomorphism, this intersection of kernels is independent of the ear decomposition; indeed, we have the following result.

LEMMA 2.5. *Let $Q = (s, t : E \to V)$ be a strongly connected quiver with ear decomposition $Q_\bullet$. For any representation $\mathbf{A}_\bullet$ of Q, there is an isomorphism*

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq K(Q_\bullet; \mathbf{A}_\bullet)$$

between the space of sections of $\mathbf{A}_\bullet$ over Q and the intersection of the kernels from (4).

PROOF. Let T be the arborescence induced by $Q_\bullet$ and r its root vertex. Using Proposition 1.6, vectors in $\mathbf{A}_r$ correspond bijectively with sections in $\Gamma(T; \mathbf{A}_\bullet)$ via the assignment that sends each x in $\mathbf{A}_r$ to the section given by

$$v \mapsto \gamma_v = \mathbf{A}_{p[v]}(x).$$

The subspace $\Gamma(Q; \mathbf{A}_\bullet) \subset \Gamma(T; \mathbf{A}_\bullet)$, is obtained by additionally enforcing compatibility across the edges in E_{ter}. Let ϵ be a terminal edge and x_r a vector in $\mathbf{A}_r$. Now the section $v \mapsto \mathbf{A}_{p[v]}(x_r)$ of $\Gamma(T; \mathbf{A}_\bullet)$ satisfies the compatibility requirement $\mathbf{A}_\epsilon(x_{s(\epsilon)}) = x_{t(\epsilon)}$ across ϵ if and only if x_r lies in the kernel of the map Δ_ϵ from (3). Thus, our x_r-induced section is compatible across all the terminal edges if and only if x_r lies in $K(Q_\bullet; \mathbf{A}_\bullet)$. □

We may safely combine this result with Proposition 1.6 to reduce a strongly connected quiver to an arborescence while preserving the space of sections.

COROLLARY 2.6. *Assuming the hypotheses of Lemma 2.5, let T be the arborescence induced by $Q_\bullet$ and r its root vertex. Let $\mathbf{A}'_\bullet$ be the representation of T prescribed by the following assignments to* *vertices $v \in V$ and non-terminal edges $e \in E - E_{\text{ter}}$:*

$$\mathbf{A}'_v := \begin{cases} \mathbf{A}_v & v \neq r, \\ K(Q_\bullet; \mathbf{A}_\bullet) & v = r; \end{cases} \quad \text{and} \quad \mathbf{A}'_e := \begin{cases} \mathbf{A}_e & s(e) \neq r, \\ \mathbf{A}_e\big|_{K(Q_\bullet;\mathbf{A}_\bullet)} & s(e) = r. \end{cases}$$

Then, there is an isomorphism of sections

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(T; \mathbf{A}'_\bullet).$$

We will use Corollary 2.6 to perform section-preserving simplifications of arbitrary (i.e., not necessarily strongly connected) quivers.

3. The Acyclic Reduction

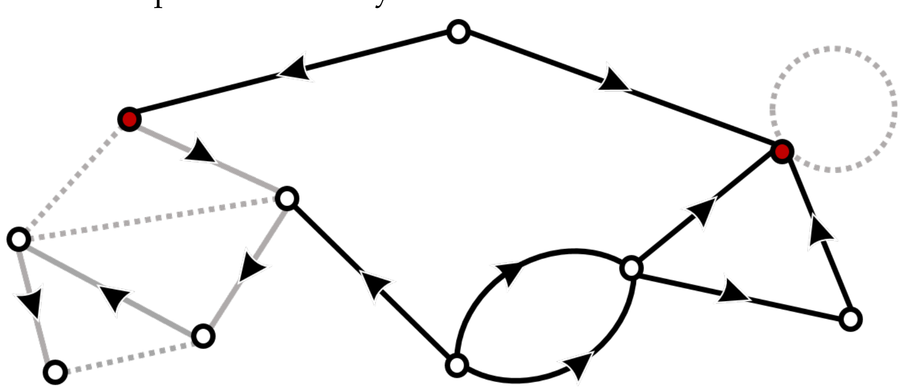
Fix a quiver $Q = (s, t : E \to V)$. A strongly connected subquiver $R \subset Q$ is *maximal* if it is not contained in a strictly larger strongly connected subquiver of Q. We denote the set of all maximal strongly connected subquivers of Q by $\mathbf{MSC}(Q)$. This set can be extracted from Q very efficiently by employing the remarkable algorithm of Tarjan [**BJG09**, Section 5.2]. Distinct subquivers in $\mathbf{MSC}(Q)$ have disjoint vertices[1]. For each R in $\mathbf{MSC}(Q)$, fix an ear decomposition $R_\bullet$ as in Definition 2.1. We write $T(R_\bullet)$ for the arborescence induced by $R_\bullet$ as in Definition 2.4, and let $E_{\text{ter}}(R_\bullet) \subset E$ be the set of terminal edges of $R_\bullet$.

DEFINITION 3.1. The **acyclic reduction** Q^* of $Q = (s, t : E \to V)$ with respect to the ear decompositions $\{R_\bullet \mid R \in \mathbf{MSC}(Q)\}$ is the subquiver $Q^* \subset Q$ defined as follows: it has the same vertex set V, while its edge set $E^* \subset E$ is given by removing all terminal edges, i.e.,

$$E^* = E - \bigcup_R E_{\text{ter}}(R_\bullet),$$

where R ranges over $\mathbf{MSC}(Q)$.

We note that the quiver Q^* is indeed acyclic (as suggested by its name) as follows. Each cycle in Q is strongly connected, hence lies in a single maximal strongly connected component $R \in \mathbf{MSC}(Q)$. But the removal of all the terminal edges $E_{\text{ter}}(R_\bullet)$ turns R into the arborescence $T(R_\bullet)$, which cannot contain any cycles. Depicted below is the quiver from the Introduction; the light-shaded edges lie within strongly-connected subquivers, whose root vertices are coloured red. The dotted edges are terminal for the associated ear decompositions, and their removal produces the acyclic reduction:



[1]If $R \neq R'$ in $\mathbf{MSC}(Q)$ share a vertex v, then there is a path (passing through v) in the union $R \cup R'$ from any vertex of R to any vertex of R' and vice versa. The existence of such paths makes $R \cup R'$ strongly connected, contradicting the maximality of either R or R'.

Our next goal is to reduce a given representation $\mathbf{A}_\bullet$ of Q to a new representation $\mathbf{A}^*_\bullet$ of Q^* in a manner that preserves the space of sections. Let $\rho : \mathbf{MSC}(Q) \to V$ be the injective *root map*, which sends each maximal strongly connected subquiver $R \subset Q$ to the root vertex of $T(R_\bullet)$. We associate to each vertex $v \in V$ the subspace $\mathbf{A}^\circ_v \subset \mathbf{A}_v$ given by

$$\mathbf{A}^\circ_v := \begin{cases} \mathbf{A}^R_v & \text{if } v = \rho(R) \text{ for some } R \in \mathbf{MSC}(Q), \\ \mathbf{A}_v & \text{otherwise,} \end{cases}$$

where, for each $R \in \mathbf{MSC}(Q)$, we write $\mathbf{A}^R_\bullet$ for the representation of $T(R_\bullet)$ described in Corollary 2.6.

DEFINITION 3.2. For each vertex $v \in V$ and strongly connected $R \in \mathbf{MSC}(Q)$, let $P^*_{v\to R}$ be the set of all paths in Q^* with source v and target $\rho(R)$. The R**-constrained space at** v is the subspace $\Lambda_{v,R} \subset \mathbf{A}^\circ_v$ given by

$$\Lambda_{v,R} := \left\{ x \in \mathbf{A}^\circ_v \mid \mathbf{A}_p(x) \in \mathbf{A}^R_{\rho(R)} \text{ for all } p \in P^*_{v\to R} \right\},$$

with the implicit understanding that $\Lambda_{v,R}$ equals $\mathbf{A}^\circ_v$ whenever $P^*_{v\to R}$ is empty.

Our next result shows that R-constrained subspaces behave well under the linear maps assigned by $\mathbf{A}_\bullet$ to edges of Q^*.

PROPOSITION 3.3. *For any edge e in E^* and subquiver $R \in \mathbf{MSC}(Q)$, the linear map $\mathbf{A}_e : \mathbf{A}_{s(e)} \to \mathbf{A}_{t(e)}$ sends $\Lambda_{s(e),R}$ to $\Lambda_{t(e),R}$.*

PROOF. Let $p = (e_1, \ldots, e_k)$ be any path in $P^*_{t(e)\to R}$ and note that the augmented path $p' = (e, e_1, \ldots, e_k)$ is an element of $P^*_{s(e)\to R}$. Now for any x in $\Lambda_{s(e),R}$ we know that $\mathbf{A}_{p'}(x)$ lies in $\mathbf{A}^R_{\rho(R)}$ by Definition 3.2. But $\mathbf{A}_{p'}(x)$ is $\mathbf{A}_p \circ \mathbf{A}_e(x)$, whence $\mathbf{A}_e(x)$ lies in $\Lambda_{t(e),R}$. □

Consider the intersection of all the R-constrained spaces at a given vertex $v \in V$, i.e, define the subspace $\Lambda_v \subset \mathbf{A}_v$ as

$$\Lambda_v := \bigcap_R \Lambda_{v,R} \tag{5}$$

where R ranges over $\mathbf{MSC}(Q)$. It follows immediately from Proposition 3.3 that for each edge e in E^* the map $\mathbf{A}_e$ sends $\Lambda_{s(e)}$ to $\Lambda_{t(e)}$.

DEFINITION 3.4. Let Q^* be the acyclic reduction of Q with respect to a choice of ear decompositions $\{R_\bullet \mid R \in \mathbf{MSC}(Q)\}$. The **acyclification** of a representation $\mathbf{A}_\bullet$ of Q is a new representation $\mathbf{A}^*_\bullet$ of Q^* which assigns to every vertex v in V the vector space

$$\mathbf{A}^*_v = \Lambda_v$$

and to every edge e in E^* the restriction of $\mathbf{A}_e$ to $\Lambda_{s(e)}$, denoted $\mathbf{A}^*_e : \Lambda_{s(e)} \to \Lambda_{t(e)}$.

As promised, our new representation $\mathbf{A}^*_\bullet$ retains full knowledge of the sections of the original representation $\mathbf{A}_\bullet$ even though it is only defined on the acyclic reduction Q^*.

PROPOSITION 3.5. *Let $\mathbf{A}_\bullet$ be a representation of a quiver $Q = (s, t : E \to V)$. Writing Q^* for the acyclic reduction of Q with respect to some choice of ear decompositions $\{R_\bullet \mid R \in \mathbf{MSC}(Q)\}$ and $\mathbf{A}^*_\bullet$ for the corresponding acyclification of $\mathbf{A}_\bullet$, there is an isomorphism of sections*

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(Q^*; \mathbf{A}^*_\bullet).$$

PROOF. First we show that a section γ in $\Gamma(Q;\mathbf{A}_\bullet)$ gives a section in $\Gamma(Q^*;\mathbf{A}^*_\bullet)$. Since $E^* \subset E$ by Definition 3.1, it suffices to prove that γ_v lies in the subspace Λ_v of $\mathbf{A}_v$ for all vertices v in V. Since γ restricts to a section in $\Gamma(R;\mathbf{A}_\bullet)$ for every subquiver $R \in \mathbf{MSC}(Q)$, it follows from Corollary 2.6 that $\gamma_{\rho(R)}$ lies in the subspace $\mathbf{A}^R_{\rho(R)}$ of $\mathbf{A}_{\rho(R)}$. Thus, for any vertex $v \in V$ and every path p in $P^*_{v\to R}$, compatibility forces $\mathbf{A}_p(\gamma_v) \in \mathbf{A}^R_{\rho(R)}$. Thus, γ_v must lie in the subspace Λ_v from (5). Now consider any edge $e \in E^*$ and note that $\mathbf{A}^*_e$ is defined simply by restricting $\mathbf{A}_e$ to the subspace $\Lambda_{s(e)}$. Thus, we obtain

$$\mathbf{A}^*_e(\gamma_{s(e)}) = \mathbf{A}_e(\gamma_{s(e)}) = \gamma_{t(e)}$$

for each such edge, and it follows that γ is a section in $\Gamma(Q^*;\mathbf{A}^*_\bullet)$. Conversely, consider a section γ^* in $\Gamma(Q^*;\mathbf{A}^*_\bullet)$. The $\mathbf{A}_\bullet$-compatibility of γ^* across every edge $e \in E^*$ follows from the fact that $\mathbf{A}^*_e$ is the restriction of $\mathbf{A}_e$; it therefore suffices to show that γ^* is also $\mathbf{A}_\bullet$-compatible across all the edges in $E - E^*$. By Definition 3.1, any such edge ϵ lies in $E_{\text{ter}}(R_\bullet)$ for a unique $R \in \mathbf{MSC}(Q)$. We know that $\Lambda_{\rho(R)}$ is a subspace of $\mathbf{A}^R_{\rho(R)}$, by (5) combined with Definition 3.2. Thus, Corollary 2.6 guarantees that γ^* is also $\mathbf{A}_\bullet$-compatible across ϵ, as desired. □

4. The Arboreal Replacement

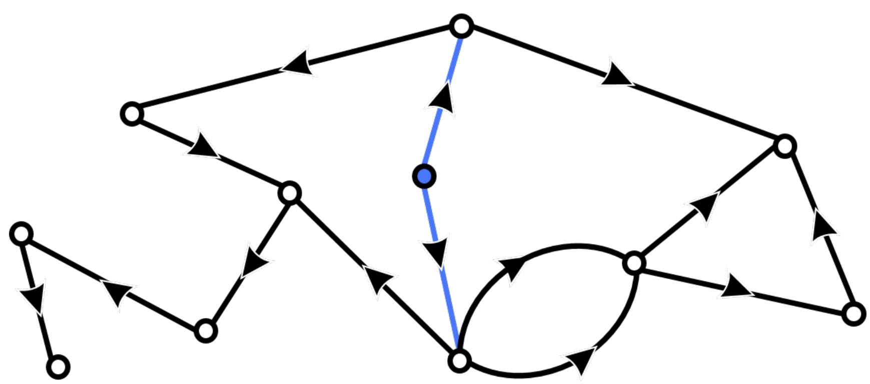
We assume here that $Q = (s,t : E \to V)$ is an acyclic quiver, so its vertex set V is partially ordered by (the reflexive closure of) the binary relation

$$u < v \text{ iff there is a path } p \text{ in } Q \text{ with } s(p) = u \text{ and } t(p) = v.$$

Let $V_{\min} \subset V$ be set of all minimal vertices with respect to this partial order — thus, a vertex v lies in $V_{\min}$ if and only if there is no edge $e \in E$ with $t(e) = v$. We fix a representation $\mathbf{A}_\bullet$ of Q, and seek to compute the space of sections $\Gamma(Q;\mathbf{A}_\bullet)$. For this purpose, it will be convenient to formally add a new vertex to Q that serves as the global minimum for the partial order described above.

DEFINITION 4.1. The **augmented quiver** Q^+ has vertices $V^+ := V \cup \{r\}$, where r is a new vertex. Its edge set E^+ is $E \cup \{e_v \mid v \in V_{\min}\}$; the sources and targets of edges in E are inherited from Q, while each new edge e_v has source r and target v in $V_{\min}$.

Drawn below is the augmented quiver corresponding to the acyclic reduction from the previous section; the root and (two) new edges e_v for the vertices $v \in V_{\min}$ are highlighted in blue.



A representation $\mathbf{A}_\bullet$ extends to Q^+ if we define

$$\mathbf{A}_r := \prod_{v \in V_{\min}} \mathbf{A}_v,$$

and let $\mathbf{A}_{e_v} : \mathbf{A}_r \to \mathbf{A}_v$ be the canonical projection map. Now each section of $\mathbf{A}_\bullet$ over Q extends uniquely to a section over Q^+, whence we have an isomorphism

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(Q^+; \mathbf{A}_\bullet). \tag{6}$$

Thus, there is no loss of generality encountered when computing the sections of $\mathbf{A}_\bullet$ over Q^+ rather than Q. We will also make frequent use of the following notion.

DEFINITION 4.2. Let $n \geq 1$ be a natural number and X, Y a pair of vector spaces. The **equaliser** of a collection of n linear maps $\{f_i : X \to Y \mid 1 \leq i \leq n\}$ is the largest subspace $\mathrm{Eq}\{f_\bullet\} \subset X$ satisfying $f_i(x) = f_j(x)$ for all x in $\mathrm{Eq}\{f_\bullet\}$ and all i, j in $\{1, \ldots, n\}$.

In practice, for finite-dimensional X the equaliser $\mathrm{Eq}\{f_\bullet\}$ can be computed by intersecting kernels of successive differences:

$$\mathrm{Eq}\{f_\bullet\} = \bigcap_{i=1}^{n-1} \ker\left(f_i - f_{i+1}\right),$$

with the understanding that for $n = 1$ this intersection over the empty set equals all of X.

DEFINITION 4.3. Assign to each vertex $v \in V^+$ a subspace $\Phi_v \subset \mathbf{A}_r$ and a linear map $\varphi_v : \Phi_v \to \mathbf{A}_v$, called the **flow space** and **flow map** of $\mathbf{A}_\bullet$ at v, inductively over the partial order $\leq$ as follows:

(1) for $v = r$, the flow space Φ_r equals $\mathbf{A}_r$, and the flow map $\varphi_r : \Phi_r \to \mathbf{A}_r$ is the identity;
(2) for $v \neq r$, let $E_{\text{in}}(v) \subset E^+$ be the (necessarily nonempty) set of all edges e satisfying $t(e) = v$. Noting that $s(e) < v$ for any such e, define the subspace $\Phi'_v \subset \mathbf{A}_r$ via

$$\Phi'_v := \bigcap_e \Phi_{s(e)},$$

where e ranges over $E_{\text{in}}(v)$. For each such e, the composition $\mathbf{A}_e \circ \varphi_{s(e)}$ restricts to a linear map $\Phi'_v \to \mathbf{A}_u$. The flow space at v is the equaliser

$$\Phi_v := \mathrm{Eq}\Big\{\mathbf{A}_e \circ \varphi_{s(e)} : \Phi'_v \to \mathbf{A}_v \mid e \in E_{\text{in}}(v)\Big\}.$$

The flow map $\varphi_v : \Phi_v \to \mathbf{A}_v$ is given by $\mathbf{A}_e \circ \varphi_{s(e)}$ for any e in $E_{\text{in}}(v)$.

By construction, the flow space Φ_v for a vertex $v \neq r$ forms a subspace of the intersection $\bigcap_u \Phi_u$ of flow spaces ranging over all preceding vertices $u < v$. Thus, we can restrict the flow map φ_u at u to a vector in the flow space Φ_v whenever $u \leq v$. Our affinity for flow spaces and maps stems mainly from the following result.

PROPOSITION 4.4. *For each vertex $v \in V^+$, let $Q^+_{\leq v}$ be the subquiver of Q^+ generated by all vertices $u \leq v$ and the edges between them. Then, γ is a section in $\Gamma(Q^+_{\leq v}; \mathbf{A}_\bullet)$ if and only if the vector $\gamma_r \in \mathbf{A}_r$ lies in the flow space Φ_v.*

PROOF. For $v = r$ the result holds because in this case the spaces below are all equal:

$$\Phi_r = \Gamma(Q^+_{\leq r}; \mathbf{A}_\bullet) = \mathbf{A}_r,$$

with the flow map $\varphi_r : \Phi_r \to \mathbf{A}_r$ being the identity. Proceeding inductively over the partial order $\leq$, consider any $v \neq r$ and assume that the desired result holds for all preceding vertices $u < v$. We must show that any $x \in \Phi_v$ generates a section in $\Gamma(Q^+_{\leq v}; \mathbf{A}_\bullet)$ via the assignment $u \mapsto \varphi_u(x)$ for every $u \leq v$. Compatibility for all edges e with $t(e) \neq v$ follows from the induction hypothesis, so it suffices to examine all edges $e \in E_{\text{in}}(v)$. For any such edge, Definition 4.3 yields

$$\mathbf{A}_e \circ \varphi_{s(e)}(x) = \varphi_v(x),$$

hence establishing the desired compatibility. Conversely, if γ is a section in $\Gamma(Q^+_{\leq v}; \mathbf{A}_\bullet)$ then it suffices to show that the vector $\gamma_r \in \mathbf{A}_r$ lies in the subspace Φ_v. By the inductive hypothesis, we have

$$\gamma_r \in \Phi'_v = \bigcap_e \Phi_{s(e)},$$

where e ranges over the edges in $E_{\text{in}}(v)$. By compatibility of γ across any such e, we have

$$\mathbf{A}_e \circ \varphi_{s(e)}(\gamma_r) = \gamma_v,$$

so γ_r lies in the equaliser $\Phi_v = \text{Eq}\left\{\mathbf{A}_e \circ \varphi_{s(e)} \mid e \in E_{\text{in}}(v)\right\}$ as desired. □

Using the fact that the quiver Q^+ is the union of the subquivers $\{Q^+_{\leq v} \mid v \in V^+\}$, we are able to describe the sections of $\mathbf{A}_\bullet$ as intersections of its flow spaces. We write $V_{\max} \subset V$ for the $\leq$-maximal vertices (i.e., the vertices which do not serve as sources of edges in E^+).

PROPOSITION 4.5. *Let $\mathbf{A}_\bullet$ be a representation of an acyclic quiver Q, and Q^+ the augmented quiver (as in Definition 4.1). We have an isomorphism*

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \bigcap_{v \in V_{\max}} \Phi_v$$

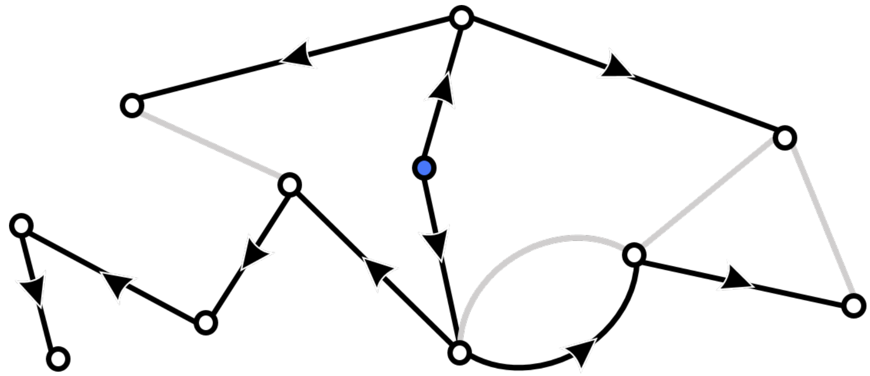
between the sections of $\mathbf{A}_\bullet$ over Q and the intersection of the flow spaces of $\mathbf{A}_\bullet$ at the maximal vertices.

PROOF. Combining (6) with Proposition 4.4 and the fact that $Q^+ = \bigcup_{v \in V} Q^+_{\leq v}$ gives

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \bigcap_{v \in V} \Phi_v.$$

Since maximal vertices have the smallest flow spaces, by Definition 4.3, the desired result follows. □

For brevity, we write $\Phi(\mathbf{A}_\bullet)$ to indicate the intersection $\bigcap_v \Phi_v$ of flow spaces ranging over $V_{\max}$ (or, equivalently, over V). By employing depth-first search [**BJG09**, Chapter 1.9] on Q^+ starting at r, one can construct a spanning arborescence $T^+ \subset Q^+$ with root r. This arborescence T^+ must necessarily contain all the vertices in V^+ and all the edges in $(E^+ - E)$, but in general it is not uniquely determined otherwise. One possible spanning arborescence for the augmented quiver drawn above is obtained by removing the light-shaded edges below:



DEFINITION 4.6. Let $T^+ \subset Q^+$ be any spanning arborescence with root r. An **arboreal replacement** of $\mathbf{A}_\bullet$ is the representation $\mathbf{A}_\bullet^+$ of T^+ that assigns

$$\mathbf{A}_v^+ := \begin{cases} \mathbf{A}_v & v \neq r, \\ \Phi(\mathbf{A}_\bullet) & v = r; \end{cases} \quad \text{and} \quad \mathbf{A}_e^+ := \begin{cases} \mathbf{A}_e & s(e) \neq r, \\ \mathbf{A}_e\big|_{\Phi(\mathbf{A}_\bullet)} & s(e) = r. \end{cases}$$

The following result is obtained by combining Proposition 4.5 with Proposition 1.6.

COROLLARY 4.7. *Let $\mathbf{A}_\bullet$ be a representation of an acyclic quiver Q and T^+ a spanning arborescence of the augmented quiver Q^+. There is an isomorphism*

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(T^+; \mathbf{A}_\bullet^+)$$

between the sections of $\mathbf{A}_\bullet$ and those of its arboreal replacement $\mathbf{A}_\bullet^+$ defined on T^+.

5. The Space of Sections

We are now ready to establish Theorem (A) from the Introduction.

THEOREM 5.1. *For any representation $\mathbf{A}_\bullet$ of a quiver $Q = (s, t : E \to V)$, the following spaces are all isomorphic:*

$$\Gamma(Q; \mathbf{A}_\bullet) \simeq \Gamma(Q^*; \mathbf{A}_\bullet^*) \simeq \Gamma(T^+; \mathbf{A}_\bullet^+) \simeq \mathbf{A}_r^+. \tag{7}$$

Here, Q^* is the acyclic reduction of Q with $\mathbf{A}_\bullet^*$ the acyclification of $\mathbf{A}_\bullet$. Similarly, writing Q^+ for the augmented quiver associated to Q^* with root r, the representation $\mathbf{A}_\bullet^+$ is the arboreal replacement of $\mathbf{A}_\bullet^*$ defined on any spanning arborescence $T^+ \subset Q^+$.

PROOF. The first isomorphism follows from Proposition 3.5, the second from Corollary 4.7, and the third from Proposition 1.6. □

Theorem (A) asserts the existence of an isomorphism $\mathbf{A}_r^+ \simeq \Gamma(Q; \mathbf{A}_\bullet)$ as a map F, which we now describe. Assuming the hypotheses and notation of Theorem 5.1, there are containments

$$\mathbf{A}_v^+ \subset \mathbf{A}_v^* \subset \mathbf{A}_v,$$

for each vertex v in V, by Definition 3.4 and Definition 4.6. Since T^+ is an arborescence, it admits a unique path $p[v]$ from its root r to any such v. This path carries a linear map $\mathbf{A}_{p[v]}^+ : \mathbf{A}_r^+ \to \mathbf{A}_v^+$, and the collection of all such linear maps (indexed over $v \in V$) assembles to furnish a single map to the direct product:

$$\mathbf{A}_r^+ \to \prod_v \mathbf{A}_v^+.$$

In light of the containments $\mathbf{A}_v^+ \subset \mathbf{A}_v$ described above, the codomain is a subspace of $\mathrm{Tot}(\mathbf{A}_\bullet)$. Thus, we obtain a linear map

$$F : \mathbf{A}_r^+ \to \mathrm{Tot}(\mathbf{A}_\bullet), \tag{8}$$

whose image of F inside $\mathrm{Tot}(\mathbf{A}_\bullet)$ is an isomorphically embedded copy of $\Gamma(Q; \mathbf{A}_\bullet)$. Although various choices (of ear decompositions and spanning arborescences) made above will produce different F's, the image of F remains invariant.

As stated in the Introduction, we will define principal components along $\mathbf{A}_\bullet$ as solutions to an optimisation problem over $\Gamma(Q; \mathbf{A}_\bullet)$. In order for this to be a non-trivial problem, one requires the dimension $d := \dim \Gamma(Q; \mathbf{A}_\bullet)$ to exceed zero. We therefore take a brief detour here in order to highlight some sufficient conditions (on Q and $\mathbf{A}_\bullet$) which give lower bounds on d. Among the simplest cases to analyse in terms of the topology of Q are the extreme ones, as recorded in the following observation.

PROPOSITION 5.2. *Let $\mathbf{A}_\bullet$ be a representation of a quiver Q.*

(1) *if Q is an arborescence with root r, then $d = \dim \mathbf{A}_r$; and*
(2) *if Q is strongly connected, then $d = 0$ for all sufficiently generic $\mathbf{A}_\bullet$.*

PROOF. The first assertion follows directly from Proposition 1.6, so we concentrate on the second assertion. Let $Q_\bullet$ be an ear decomposition of Q (see Definition 2.1) and r a vertex in Q_1. By strong connectedness, there exists a path p in Q from r to itself, which carries an endomorphism $\mathbf{A}_p : \mathbf{A}_r \to \mathbf{A}_r$. Now any section γ of Q must satisfy $\mathbf{A}_p(\gamma_r) = \gamma_r$. For generic $\mathbf{A}_\bullet$, this endomorphism $\mathbf{A}_p$ will not have 1 as an eigenvalue, so γ_r must be zero. The result now follows from applying Proposition 1.6 to the arborescence induced by $Q_\bullet$. □

Although the result in part (2) of Proposition 5.2 might appear disappointing at first glance, we note that there are several interesting non-generic families of linear maps which do admit 1 as an eigenvalue, such as those arising from row-stochastic matrices. Moreover, general quivers are neither strongly connected nor arboreal but lie somewhere in between. Using Proposition 3.5, any representation of an arbitrary quiver can be reduced to a representation of an acyclic quiver while preserving d, so it remains to provide lower bounds on d for representations of acyclic quivers.

PROPOSITION 5.3. *Let Q be an acyclic quiver with minimal vertices $V_{\min}$ and maximal vertices $V_{\max}$. For any representation $\mathbf{A}_\bullet$ of Q, we have*

$$\dim \Gamma(Q; \mathbf{A}_\bullet) \geq \sum_{u \in V_{\min}} \dim \mathbf{A}_u - \sum_{v \in V_{\max}} (n_v - 1) \dim \mathbf{A}_v,$$

where n_v is the total number of paths in the augmented quiver Q^+ from the root r to the vertex v.

PROOF. By Definition 4.3, the flow space $\Phi_r = \mathbf{A}_r$ has dimension $\sum_{u \in V_{\min}} \dim \mathbf{A}_u$. We claim that the flow space Φ_v at a vertex $v \in V_{\max}$ has codimension at most $(n_v - 1) \dim \mathbf{A}_v$ in Φ_r. To establish this claim, let $\{f_k : \mathbf{A}_r \to \mathbf{A}_v \mid 1 \leq k \leq n_v\}$ be the linear maps carried by paths from r to v, and examine the $(n_v - 1)$ kernels of the successive differences

$$\Delta_k = (f_k - f_{k+1}).$$

Since each $\ker(\Delta_k)$ has codimension at most $\dim \mathbf{A}_v$ in Φ_r, and since the codimension of their intersection is at most the sum of these codimensions, we have $\mathrm{codim} \Phi_v \leq (n_v - 1) \dim \mathbf{A}_v$ as claimed. The inequality in the statement now follows from Proposition 4.5. □

6. Principal Components via Optimisation

Here we will define principal components with respect to a quiver representation as solutions to an optimisation problem over the space of sections. To this end, let us first recall the starting point, ordinary principal components analysis (PCA).

DEFINITION 6.1. Let $D := \{y_1, \ldots, y_m\}$ be a finite collection of mean-centred[2] vectors in $\mathbb{R}^n$; the *sample covariance* of D is the $n \times n$ symmetric matrix

$$S := \frac{1}{m}\sum_{i=1}^{m} y_i y_i^{\mathsf{T}},$$

where $^{\mathsf{T}}$ indicates transpose. Assuming that the top r eigenvalues $\lambda_1 > \cdots > \lambda_r$ of S are distinct, the r-th **principal component** $\mathbf{PC}_r(D)$ of D is the λ_r-eigenspace of S.

Since the r-th principal component is a one-dimensional subspace of $\mathbb{R}^n$, it is standard practice to represent it by any constituent nonzero vector in $\mathbf{PC}_r(D)$. Treating the sample covariance matrix as a bilinear form on $\mathbb{R}^n$ allows us to interpret principal components in terms of the following variance maximisation problem:

$$\max_X \operatorname{tr}(X^{\mathsf{T}} S X) \text{ subject to } X^{\mathsf{T}} X = \mathrm{id}_r. \tag{9}$$

Here tr indicates trace and id_r is the $r \times r$ identity matrix. The columns of an optimal $n \times r$ matrix X form an orthonormal basis for the space $\mathbf{PC}_{\leq r}(D)$ spanned by the top r principal components, and solving (9) for increasing r gives the individual principal components in descending order.

6.1. Principal components along quiver representations. Consider a quiver Q and fix a representation $\mathbf{A}_\bullet$ of Q valued in real vector spaces. Henceforth we will fix an isomorphism $\mathbb{R}^{\dim \mathbf{A}_v} \xrightarrow{\simeq} \mathbf{A}_v$ for each vertex v in V, which allows us to impose (once and for all) an inner product structure on each $\mathbf{A}_v$. Writing n for the dimension of $\mathrm{Tot}(\mathbf{A}_\bullet)$,

$$n = \sum_{v \in V} \dim \mathbf{A}_v,$$

we inherit an isomorphism $\mathbb{R}^n \xrightarrow{\simeq} \mathrm{Tot}(\mathbf{A}_\bullet)$ and a concomitant inner product structure on the total space of $\mathbf{A}_\bullet$. Making choices of ear decompositions and spanning arborescences for Q produces a map $F : \mathbb{R}^d \to \mathrm{Tot}(\mathbf{A}_\bullet)$, described in (8), where $d = \dim \Gamma(Q; \mathbf{A}_\bullet)$. Expressed in terms of the chosen isomorphisms, F becomes a full-rank $n \times d$ matrix whose image is an embedded copy of $\Gamma(Q; \mathbf{A}_\bullet)$ inside $\mathbb{R}^n$. We are therefore able to define principal components relative to this embedding F.

DEFINITION 6.2. Given any mean-centred finite subset D of $\mathbb{R}^n \simeq \mathrm{Tot}(\mathbf{A}_\bullet)$, let S be the sample covariance (as in Definition 6.1). For each $r \leq d$, consider the optimisation problem over all $n \times r$ matrices $X = [x_1\ x_2\ \cdots\ x_r]$ prescribed by

$$\max_X \operatorname{tr}(X^{\mathsf{T}} S X) \quad \text{subject to} \quad X^{\mathsf{T}} X = \mathrm{id}_r \text{ and } x_1, \ldots, x_r \in \Gamma(Q; \mathbf{A}_\bullet). \tag{10}$$

The **space of top r principal components** of D along $\mathbf{A}_\bullet$ is the subspace $\mathbf{PC}_{\leq r}(D; \mathbf{A}_\bullet)$ of $\mathbb{R}^n$ determined by the column span

$$\mathbf{PC}_{\leq r}(D; \mathbf{A}_\bullet) = \operatorname{span}\{x_1, \ldots, x_r\}$$

[2] i.e., $\frac{1}{m}\sum_i y_i$ lies at the origin

of an optimal matrix X.

It is possible to uniquely construct an optimal solution X_* to (10) by proceeding one column at a time and imposing the orthogonality of each column with respect to all of the preceding columns. The ***r*-th principal component** of D along $\mathbf{A}_\bullet$ is the subspace $\mathbf{PC}_r(D;\mathbf{A}_\bullet)$ spanned by the r-th column of X_*. In sharp contrast to the ordinary principal components from Definition 6.1, these principal components along $\mathbf{A}_\bullet$ need not be eigenvectors of the covariance matrix S. There are, however, two special cases where ordinary principal components coincide with their quiver-compatible avatars.

PROPOSITION 6.3. *Assume that one of the two conditions below holds:*

(1) *either $D \subset \mathbb{R}^n$ lies entirely in the subspace $\Gamma(Q;\mathbf{A}_\bullet)$, or*
(2) *the edge set of Q is empty.*

Then $\mathbf{PC}_r(D) = \mathbf{PC}_r(D;\mathbf{A}_\bullet)$ for every $r \leq d$.

PROOF. If $D \subset \Gamma = \Gamma(Q;\mathbf{A}_\bullet)$, then the sample covariance S restricts to an endomorphism of Γ. For all $r \leq d$, the columns of any matrix X that maximises (9) must also lie in Γ. Thus, such an X also maximises (10). Finally, if there are no edges in Q then Γ equals all of $\mathbb{R}^n$ so (10) reduces to (9). □

In its most general form, linearly constrained PCA can be described as follows. The space of top r principal components of $D \subset \mathbb{R}^n$ (with sample covariance S), constrained by some $n \times c$ matrix W, is the span of the columns of an optimal $n \times r$ matrix X in

$$\max_X \operatorname{tr}(X^\mathsf{T} S X) \quad \text{subject to} \quad X^\mathsf{T} X = \mathrm{id}_r \text{ and } W^\mathsf{T} X = 0.$$

This formulation follows from [**DK96**, Equation 7.4], and it is usually assumed that $W^\mathsf{T} W = \mathrm{id}_c$. Evidently, finding principal components along a quiver representation is a special instance of constrained PCA, provided we have access to an orthogonal basis for the complement of $\Gamma(Q;\mathbf{A}_\bullet)$ in $\mathrm{Tot}(\mathbf{A}_\bullet)$.

6.2. Alternate perspectives. Here we define two more optimisation problems related to (10); as before, both will require a fixed choice of embedding $F : \mathbb{R}^d \to \mathrm{Tot}(\mathbf{A}_\bullet)$ of $\Gamma(Q;\mathbf{A}_\bullet)$, where the map F is viewed as an $n \times d$ matrix. Here is the first one, which is defined over the space over $d \times r$ matrices Y:

$$\max_Y \operatorname{tr}(Y^\mathsf{T} F^\mathsf{T} S F Y) \quad \text{subject to} \quad Y^\mathsf{T}(F^\mathsf{T} F) Y = \mathrm{id}_r. \tag{11}$$

The $n \times n$ matrix $B := FF^\mathsf{T}$ serves as a (not necessarily orthogonal) projection onto the image of F. Now, we set $S_B := BSB$ and consider another optimisation problem defined over $n \times r$ matrices Z:

$$\max_Z \operatorname{tr}(Z^\mathsf{T} S_B Z) \quad \text{subject to} \quad Z^\mathsf{T}(B^2) Z = \mathrm{id}_r. \tag{12}$$

Although the r columns of Z can be any B^2-orthonormal vectors in $\mathrm{Tot}(\mathbf{A}_\bullet)$, the optimal directions will lie in $\Gamma(Q;\mathbf{A}_\bullet)$ because S_B restricts to an endomorphism of $\Gamma(Q;\mathbf{A}_\bullet)$ for any $r \leq d$. Our next result establishes the equivalence of these two alternate perspectives with the original one from Definition 6.2.

PROPOSITION 6.4. *The maximum values of the three optimisation problems* (10), (11) *and* (12) *are all the same. Moreover, a matrix X maximises* (10) *if and only if matrix Y maximises* (11) *if and only if matrix Z maximises* (12), *where*

$$X = FY = BZ.$$

PROOF. We first show that Z maximises (12) if and only if BZ maximizes (10). Since $\Gamma = \Gamma(Q; \mathbf{A}_\bullet)$ is the image of B, it follows that the columns of BZ all lie in Γ. Moreover, we have $(BZ)^\mathsf{T}(BZ) = Z^\mathsf{T}(B^2)Z = \mathrm{id}_r$, so BZ is orthogonal and satisfies all the constraints of (10). Moreover, we have

$$\begin{aligned}\mathrm{tr}(Z^\mathsf{T} S_B Z) &= \mathrm{tr}(Z^\mathsf{T} \cdot BSB \cdot Z) \\ &= \mathrm{tr}\left((BZ)^\mathsf{T} S (BZ)\right).\end{aligned}$$

Conversely, given some X maximising (10), its columns x_i are orthonormal vectors in Γ, hence $x_i = Bz_i$ for some $z_i \in \mathrm{Tot}(\mathbf{A}_\bullet)$. Letting Z be the matrix of columns z_i gives a solution to (12) with the same maximal value (as confirmed by the trace calculation above). This gives the desired equivalence of (10) and (12). Turning now to (11), assume again that Z maximises (12) and let $Y = F^\mathsf{T} Z$, so

$$(FY)^\mathsf{T}(FY) = (BZ)^\mathsf{T}(BZ) = \mathrm{id}_r.$$

Computing the relevant trace for (11) gives

$$\begin{aligned}\mathrm{tr}(Y^\mathsf{T} F^\mathsf{T} S F Y) &= \mathrm{tr}(Z^\mathsf{T} F F^\mathsf{T} S F F^\mathsf{T} Z) \\ &= \mathrm{tr}(Z^\mathsf{T} BSBZ) \\ &= \mathrm{tr}(Z^\mathsf{T} S_B Z).\end{aligned}$$

Thus, the value of the objective function of (12) at Z equals the value of the objective function of (11) at $Y = F^\mathsf{T} Z$. Conversely, given some Y maximising (11), its image FY is a matrix of orthogonal vectors in Γ, hence lies in the feasible set for (10), with the trace of $X = FU = BV$ in (10) being the same as the trace of Y in (11). □

We consider (10) an **implicit** version of the optimisation problem to determine principal components along quiver representations, while (11) and (12) are its **parametrised** and **projected** variants. Thanks to the preceding result, it becomes possible to freely translate between these three perspectives. In practice, the dimension d of $\Gamma(Q; \mathbf{A}_\bullet)$ is much smaller than the ambient dimension n of $\mathrm{Tot}(\mathbf{A}_\bullet)$, so one might wish to work with the optimisation problem (11) in this smaller space. An algorithmic approach to (12) that similarly reduces to a smaller space has been studied in **[Gol73]**.

REMARK 6.5. The argument invoked in the proof of Proposition 6.4 simplifies considerably if the $n \times d$ matrix F has orthonormal columns. In this case, the matrices Y in (11) satisfy $Y^\mathsf{T} Y = \mathrm{id}_r$. Moreover, the matrix Z that maximises (12) satisfies $Z^\mathsf{T} Z = \mathrm{id}_r$. This is because $B = FF^\mathsf{T}$ is an orthogonal projection onto Γ, so $v \in \Gamma$ if and only if $Bv = v$. Since the columns of Z are in Γ at the optimum, we have $BZ = Z$ and hence $\mathrm{id}_r = Z^\mathsf{T} B^\mathsf{T} BZ = Z^\mathsf{T} Z$, as claimed.

7. Principal Components as Generalised Eigenvectors

We have already noted that – aside from some very special cases as in Proposition 6.3 – the principal components $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ of Definition 6.2 are not eigenvectors of the sample covariance S. Here we remedy this defect by providing a spectral interpretation for $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$. All scalars, vectors and matrices described below live over the field of real numbers.

DEFINITION 7.1. Fix two identically-sized square matrices A and B. The **generalised eigenvalues** of the matrix pencil $A - \lambda B$ are the solutions λ to $\det(A - \lambda B) = 0$. We call a non-zero vector x with $Ax = \lambda Bx$ a **generalised eigenvector**, with λ its generalised eigenvalue[3].

Our main tool in the quest to interpret quiver prinicipal components as generalised eigenvectors is the generalised singular value decomposition (GSVD) [**VL76**, Theorem 2].

THEOREM 7.2. [GSVD] *Given positive integers $a \geq b \geq c$, fix an $(a \times c)$ matrix A and a $(b \times c)$ matrix B. There exist*

(1) *orthogonal matrices W_A and W_B of size $a \times a$ and $b \times b$ respectively,*
(2) *(rectangular) diagonal matrices Δ and Σ of size $a \times c$ and $b \times c$ respectively, and*
(3) *a $c \times c$ invertible matrix G,*

satisfying both

$$A = W_A \Delta G \qquad \textit{and} \qquad B = W_B \Sigma G.$$

The matrices W_A, W_B and G are not uniquely determined, but the diagonal entries of Δ and Σ are completely specified (up to reordering) by A and B. Moreover, these diagonals are non-negative, with the number of nonzero entries coinciding with the ranks of A and B respectively. We note en passant that a different generalisation of the singular value decomposition [**VL76**, Theorem 3] also appears in the context of constrained PCA, and that a discussion of GSVD naming conventions can be found in [**TH01**, Section 5.5]. Returning to the setting of interest, we fix a representation $\mathbf{A}_\bullet$ of a quiver Q and select a full-rank $n \times d$ matrix $F : \mathbb{R}^d \to \mathrm{Tot}(\mathbf{A}_\bullet)$ whose image is $\Gamma(Q; \mathbf{A}_\bullet)$. The following result is Theorem (B) from the Introduction.

THEOREM 7.3. *Let S be the sample covariance of a sufficiently generic mean-centered subset $D = \{y_1, \ldots, y_m\} \subset \mathbb{R}^n$ of cardinality $m \geq n$. For each $r \leq d$, the r-th principal component $\mathbf{PC}_r(D; \mathbf{A}_\bullet)$ is spanned by Fu_r, where u_r is the eigenvector of the matrix pencil $F^\mathsf{T} SF - \lambda(F^\mathsf{T} F)$ corresponding to its r-th largest generalised eigenvalue.*

PROOF. Let M denote the $m \times n$ matrix whose i-th row is the normalised vector $y_i/\sqrt{m}$, so that the sample covariance satisfies $S = M^\mathsf{T} M$. Noting that $m \geq n \geq d$, we apply the GSVD from Theorem 7.2 to the $m \times d$ matrix $A = MF$ and the $n \times d$ matrix $B = F$. This produces factorisations

$$MF = W_A \Delta G \qquad \text{and} \qquad F = W_B \Sigma G$$

with orthogonal W_A, W_B, invertible G, and diagonal Δ, Σ. Since D is generic and F has full rank, we may safely assume that the diagonal entries of Δ and Σ are nonzero. And by orthogonality of both the W-matrices, we obtain two new identities

$$(MF)^\mathsf{T}(MF) = G^\mathsf{T} \Delta^2 G \qquad \text{and} \qquad F^\mathsf{T} F = G^\mathsf{T} \Sigma^2 G. \tag{13}$$

Since $S = M^\mathsf{T} M$ by design, the first identity reduces to $F^\mathsf{T} SF = G^\mathsf{T} \Delta^2 G$. Let us write $\{\delta_1, \ldots, \delta_d\}$ and $\{\sigma_1, \ldots, \sigma_d\}$ for the (necessarily nonzero) diagonal entries of Δ and Σ respectively, and denote by g_i the i-th column of G^{-1}. It follows from (13) that g_i is a generalised

[3]We hope there is no confusion with the (somewhat more standard) use of generalised eigenvector to mean a nonzero solution x of $(A - \lambda I)^m x = 0$ for $m \geq 1$.

eigenvector for the $d \times d$ matrix pencil $(F^\mathsf{T} SF) - \lambda \cdot (F^\mathsf{T} F)$, corresponding to the generalised eigenvalue $\lambda_i := \delta_i^2/\sigma_i^2$. In other words, we have

$$(F^\mathsf{T} SF)g_i = \lambda_i \cdot (F^\mathsf{T} F)g_i. \tag{14}$$

The top $d \times d$ block Σ_d of Σ is invertible because its diagonal has nonzero entries. Since G is also invertible, the product $\Sigma_d G$ permutes the set of $d \times r$ matrices via $Y \mapsto Y_\circ = \Sigma_d GY$, which allows us to re-express the optimisation (11) in a particularly convenient form. To this end, we calculate:

$$\begin{aligned}
Y^\mathsf{T}(F^\mathsf{T} SF)Y &= Y^\mathsf{T}(G^\mathsf{T}\Delta^2 G)Y && \text{by (13)} \\
&= (G^{-1}\Sigma_d^{-1}Y_\circ)^\mathsf{T}(G^\mathsf{T}\Delta^2 G)(G^{-1}\Sigma_d^{-1}Y_\circ) && \text{since } Y_\circ = \Sigma_d GY \\
&= Y_\circ^\mathsf{T}\, \Sigma_d^{-1}\Delta^2\Sigma_d^{-1}\, Y_\circ && \text{after two cancellations.}
\end{aligned}$$

Now the intermediate product $\nabla := \Sigma_d^{-1}\Delta^2\Sigma_d^{-1}$ is a $d \times d$ diagonal matrix whose i-th diagonal entry is $\lambda_i = \delta_i^2/\sigma_i^2$. Reordering basis vectors if necessary, we can assume without loss of generality that $\lambda_1 > \cdots > \lambda_d$. The change of variables $Y \mapsto Y_\circ$ transforms the optimisation problem from (11) into

$$\max_{Y_\circ} \operatorname{tr}(Y_\circ^\mathsf{T} \nabla Y_\circ) \quad \text{subject to} \quad Y_\circ^\mathsf{T} Y_\circ = \mathrm{id}_r.$$

This is the ordinary PCA optimisation (9), which admits a unique solution Y_* obtained by successively increasing r. Since ∇ is diagonal, the i-th column of Y_* is the i-th elementary basis vector. Thus, the columns $\{u_1, \ldots, u_r\}$ of $U = G^{-1}\Sigma_d^{-1}Y_*$ lie in the directions of the corresponding columns of G^{-1}. By (14), these columns are generalised eigenvectors associated to the r largest generalised eigenvalues of our matrix pencil. Finally, applying F to U gives the principal components along the quiver representation as in Proposition 6.4. □

It follows that the top principal component is Fu, where u maximises the Rayleigh quotient

$$\frac{u^\mathsf{T}(F^\mathsf{T} SF)u}{u(F^\mathsf{T} F)u}, \tag{15}$$

but in general for $r > 1$ the optimisation (11) is not equivalent to a single trace ratio problem (see [**NBS12**]).

REMARK 7.4. Since the embedding $F : \mathbb{R}^d \to \mathrm{Tot}(\mathbf{A}_\bullet)$ has rank d, the $d \times d$ matrix $F^\mathsf{T} F$ is invertible. We can therefore convert the generalised eigenproblem of Theorem 7.3 into the usual eigenvector problem $F^+ SFx = \lambda x$, where $F^+ = (F^\mathsf{T} F)^{-1}F^\mathsf{T}$ is the pseudo-inverse. However, as explained in [**Gol73**, Section 4], it is often preferable to work with the generalised eigenvalue problem as the matrix $F^+ SF$ may not be symmetric. And depending on the condition number of $(F^\mathsf{T} F)^{-1}$, the numerical stability might be worse .

EXAMPLE 7.5. Consider the quiver $u\bullet \xrightarrow{\;e\;} \bullet v$ with representation:

$$\mathbb{R}^p \bullet \xrightarrow{\;J\;} \bullet\, \mathbb{R}^q.$$

Writing $n = p + q$ for the dimension of the total space, the $n \times n$ sample covariance S of some $D \subset \mathbb{R}^n$ and the embedding $F : \mathbb{R}^p \to \mathbb{R}^n$ can be written as

$$S = \begin{bmatrix} S_{uu} & S_{uv} \\ S_{vu} & S_{vv} \end{bmatrix}, \qquad F = \begin{bmatrix} \mathrm{id}_p \\ J \end{bmatrix},$$

where $S_{vu} = S_{uv}^{\mathsf{T}}$. Theorem 7.3 shows that the principal components are given by the generalised eigenvectors of the matrix pencil $A - \lambda B$ spanned by

$$A = S_{uu} + J^{\mathsf{T}} S_{vu} + S_{uv} J + J^{\mathsf{T}} S_{vv} J, \qquad B = \mathrm{id}_p + J^{\mathsf{T}} J.$$

In the special case where D lies in the image of F, we have

$$S = \begin{bmatrix} \mathrm{id}_p \\ J \end{bmatrix} S_{uu} \begin{bmatrix} \mathrm{id}_p & J^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} S_{uu} & S_{uu} J^{\mathsf{T}} \\ J S_{uu} & J S_{uu} J^{\mathsf{T}} \end{bmatrix},$$

and the matrix pencil is spanned by

$$A = S_{uu} + J^{\mathsf{T}} J S_{uu} + S_{uu} J^{\mathsf{T}} J + J^{\mathsf{T}} J S_{uu} J^{\mathsf{T}} J, \qquad B = \mathrm{id}_p + J^{\mathsf{T}} J.$$

If, in addition, $J^{\mathsf{T}} J$ equals $\eta \mathrm{id}_p$ for some scalar η, then this specialises further to give the matrix pencil spanned by $A = (1 + 2\eta + \eta^2) S_{uu}$ and $B = (1 + \eta)\mathrm{id}_p$. Now the principal components along the quiver representation are given by $F\xi$, where ξ are the usual principal components of D restricted to the vector space $\mathbb{R}^p$ on the first vertex of the quiver.

8. Learning Quiver Representations

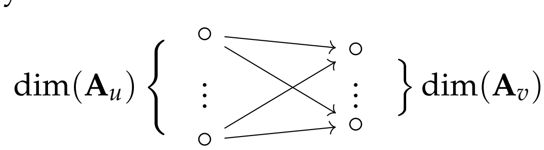
We conclude this paper with a discussion focused on the problem of learning quiver representations from observed data. Fix a quiver $Q = (s, t : E \to V)$, and assume that we have full knowledge of the real vector spaces $\{\mathbf{A}_v \mid v \in V\}$ assigned by some Q-representation $\mathbf{A}_\bullet$ to all the vertices. However, none of the linear maps $\mathbf{A}_e : \mathbf{A}_{s(e)} \to \mathbf{A}_{t(e)}$ are known. Instead, we are given access to mean-centred data $\{y_1, \ldots, y_m\}$, where each y_i is a vector in the total space $\mathrm{Tot}(\mathbf{A}_\bullet) \simeq \mathbb{R}^n$. Our task is to determine the $\mathbf{A}_e$ maps that best fit the available data; here we will show how in special cases this task reduces to well-studied problems. It will be convenient to define, for each vertex v, the $m \times \dim \mathbf{A}_v$ matrix Y_v whose i-th row is the part of y_i that lies in $\mathbf{A}_v$.

EXAMPLE 8.1. Consider the quiver $u\bullet \xrightarrow{e} \bullet v$ with representation $\mathbf{A}_u \bullet \xrightarrow{\mathbf{A}_e} \bullet \mathbf{A}_v$, with matrix $\mathbf{A}_e$ unknown. Given data $y_i = (y_{i,u}, y_{i,v}) \in \mathbf{A}_u \times \mathbf{A}_v$ for $i \in \{1, \ldots, m\}$, minimising the Euclidean distance between $y_{i,v}$ and $\mathbf{A}_e y_{i,u}$ for each i gives the least squares optimisation problem

$$\min_{\mathbf{A}_e} \| Y_v - Y_u \mathbf{A}_e^{\mathsf{T}} \|.$$

Thus, the optimal estimate for $\mathbf{A}_e^{\mathsf{T}}$ is $(Y_u)^+ Y_v$, where $(Y_u)^+$ indicates the Moore-Penrose inverse of Y_u.

The preceding example can equivalently be viewed as training (or, learning the $\dim(\mathbf{A}_u) \times \dim(\mathbf{A}_v)$ parameters in) a linear neural network with full bipartite connections between a single input and output layer:



REMARK 8.2. More generally, a linear neural network with k layers corresponds to learning a quiver representation on the quiver with k edges

$$\bullet \xrightarrow{e_1} \bullet \xrightarrow{e_2} \cdots \xrightarrow{e_{k-1}} \bullet \xrightarrow{e_k} \bullet$$

Each vertex v is replaced by $\dim(\mathbf{A}_v)$ scalar nodes, with full bipartite connections between nodes in adjacent layers. A more general architecture could involve other quivers. For example, loops arise from lateral interactions [**BH95**, Figure 5]. The setting of learning parameters in linear neural networks with two layers is itself closely connected to principal components analysis [**BH89**].

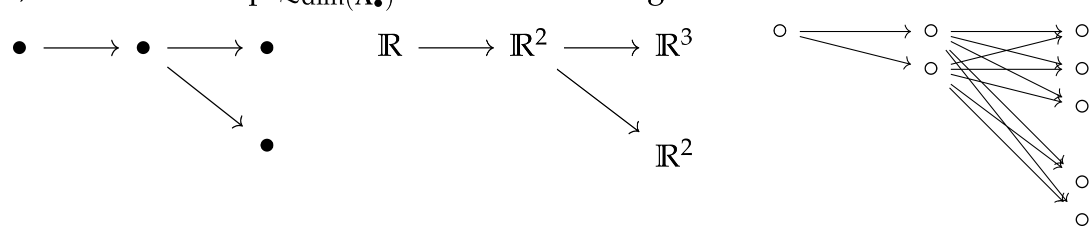
One way to extend the above to more general quivers is to learn the map on each edge e independently, which amounts to minimising the objective function:

$$\sum_{e \in E} \left(\| Y_{t(e)} - Y_{s(e)} \mathbf{A}_e^{\mathsf{T}} \|^2 \right). \tag{16}$$

Now the estimate for each edge map $\mathbf{A}_e$ is given by Example 8.1. If the quiver Q is an arborescence, then the optimisation (16) falls into the setting of a *Gaussian graphical model* [**Lau96**, **Sul18**] associated to a certain directed acyclic graph with $\dim \mathrm{Tot}(\mathbf{A}_\bullet)$ vertices, as we now describe.

DEFINITION 8.3. Let $\delta : V \to \mathbb{Z}_{\geq 0}$ be a function from the vertices of an arborescence Q to the non-negative integers. The δ**-blowup** of Q is the quiver Q_δ where each $v \in V$ is replaced by $\delta(v)$ vertices, and each edge $e \in E$ is replaced by a complete directed bipartite graph whose edges go from the $\delta(s(e))$ vertices replacing $s(e)$ to the $\delta(t(e))$ vertices replacing $t(e)$.

The directed acyclic graph of interest to us here is the δ-blowup of the arborescence Q where $\delta(v) = \dim \mathbf{A}_v$. We denote this blowup by $Q_{\dim(\mathbf{A}_\bullet)}$. For instance, if Q is the arborescence on the left and $\mathbf{A}_\bullet$ is the representation (known only on the vertices) depicted in the middle, then the blowup $Q_{\dim(\mathbf{A}_\bullet)}$ is shown to the right.



The entries of the unknown matrices $\mathbf{A}_e$ become unknown scalar weights on the edges of $Q_{\dim(\mathbf{A}_\bullet)}$. Maximum likelihood estimation in the Gaussian graphical model learns the weights on these edges by minimising least squares error. Since this is equivalent to (16), it gives an identical estimate for the unknown maps in the quiver representation.

Although Definition 8.3 extends verbatim to the case where Q is not an arborescence, the maximal likelihood estimation strategy described above is restricted to the setting of an arborescence. This is because weights of incoming edges at a vertex of the directed acyclic graph are summed over in a graphical model [**Sul18**, Equation (13.2.3)]. By comparison, in the quiver setting we do not sum incoming edges from different vertices of the quiver in (16). Thus the above strategy only works when each vertex in the quiver has at most one incoming edge.

The local assumption governing the choice of objective function in (16) is that the maps $\mathbf{A}_e$ can be learned independently of one another; this does not take into account the goodness of fit of data along longer paths in the quiver. Given such a path p, one may wish to minimise the distance between $y_{i,t(p)}$ and $\mathbf{A}_p y_{i,s(p)}$. This yields an immediate generalisation of the objective function (16), where one sums the contributions of each path in Q, rather than just over each edge. For acyclic quivers, such an optimisation can be approached by using a suitable partial order on edges, but it is more complicated for quivers with cycles. We defer a more general study of learning maps in quiver representations to future work.

Our final example is an illustration of finding principal components along a learned quiver representation. This combines parameter estimation with principal component analysis, as is also seen in [**WH09**, **MWH12**, **TA18**].

EXAMPLE 8.4. Consider once again the quiver with one edge $e : u \to v$ and representation $\mathbb{R}^p \to \mathbb{R}^q$ with unknown $\mathbf{A}_e$. The best estimate is given by $(Y_u^+ Y_v)^\intercal$, as described in Example 8.1. Thus, a parameterisation of the space of sections $\Gamma(Q; \mathbf{A}_\bullet)$ is given by

$$F = \begin{bmatrix} I \\ (Y_u^+ Y_v)^\intercal \end{bmatrix}.$$

The top principal component along the quiver representation is the direction in the image of F along which there is maximum variance in the data. This can be computed using Theorem 7.3 via the matrix pencil from Example 7.5, provided that we set $J = (Y_u^+ Y_v)^\intercal$.

References

[AJ20] Marco Antonio Armenta and Pierre-Marc Jodoin. The representation theory of neural networks. *arXiv:2007.12213 [cs.LG]*, 2020.

[AKRS20] Carlos Améndola, Kathlén Kohn, Philipp Reichenbach, and Anna Seigal. Invariant theory and scaling algorithms for maximum likelihood estimation. *arXiv preprint arXiv:2003.13662 [math.ST]*, 2020.

[ARS97] Maurice Auslander, Idun Reiten, and Sverre Smalø. *Representation theory of Artin algebras*, volume 36 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1997.

[BGP73] I N Bernšteĭn, I M Gelfand, and V A Ponomarev. Coxeter functors and Gabriel's theorem. *Russian Mathematical Surveys*, 28(2):17–32, 1973.

[BH89] Pierre Baldi and Kurt Hornik. Neural networks and principal component analysis: Learning from examples without local minima. *Neural networks*, 2(1):53–58, 1989.

[BH95] Pierre F Baldi and Kurt Hornik. Learning in linear neural networks: A survey. *IEEE Transactions on neural networks*, 6(4):837–858, 1995.

[BHS+19] Danielle A Brake, Jonathan D Hauenstein, Frank-Olaf Schreyer, Andrew J Sommese, and Michael E Stillman. Singular value decomposition of complexes. *SIAM Journal on Applied Algebra and Geometry*, 3(3):507–522, 2019.

[BJG09] Jørgen Bang-Jensen and Gregory Gutin. *Digraphs* 2nd ed. Springer, 2009.

[CGN16] Justin Curry, Robert Ghrist, and Vidit Nanda. Discrete Morse theory for computing cellular sheaf cohomology. *Foundations of Computational Mathematics*, 16(4):875–897, 2016.

[Cur13] Justin Curry. Sheaves, cosheaves and applications. *arXiv:1303.3255 [math.AT]*, 2013.

[DK96] Konstantinos I Diamantaras and Sun Yuan Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc., 1996.

[DM20] Harm Derksen and Visu Makam. Maximum likelihood estimation for matrix normal models via quiver representations. *arXiv preprint arXiv:2007.10206 [math.RT]*, 2020.

[DW11] Harm Derksen and Jerzy Weyman. The combinatorics of quiver representations. *Annales de L'Institut Fourier, Grenoble*, 61(3):1061–1131, 2011.

[DW17] H Derksen and J Weyman. *An Introduction to Quiver Representations*. Number 184 in Graduate Studies in Mathematics. The American Mathematical Society, 2017.

[FM21] Cole Franks and Visu Makam. IPCA and stability for star quivers. *in preparation*, 2021.

[Gab72] Peter Gabriel. Unzerlegbare darstellungen I. *Manuscripta Mathematica*, 6:71–103, 1972.

[Gin09] Victor Ginzburg. Lectures on Nakajima's quiver varieties. *arXiv:0905.0686 [math.RT]*, 2009.

[GMV96] Sergei Gelfand, Robert MacPherson, and Kari Vilonen. Perverse sheaves and quivers. *Duke Math Journal*, 83(3):621–643, 1996.

[Gol73] Gene H Golub. Some modified matrix eigenvalue problems. *SIAM Review*, 15(2):318–334, 1973.

[GP10] M Gross and R Pandharipande. Quivers, curves, and the tropical vertex. *Portugaliae Mathematica*, 67(2):211–259, 2010.

[Her08] Martun Herschend. Tensor products on quiver representations. *Journal of Pure and Applied Algebra*, 212(2):452–469, 2008.

[HKKP17] Fabian Haiden, Ludmil Katzarkov, Maxim Kontsevich, and Pranav Pandit. Semistability, modular lattices, and iterated logarithms. *arXiv:1706.01073 [math.RT]*, 2017.

[HMR20] Karim Halaseh, Tommi Muller, and Elina Robeva. Orthogonal decomposition of tensor trains. *arXiv preprint arXiv:2010.04202 [math.NA]*, 2020.

[HMS18] Shaun Harker, Konstantin Mischaikow, and Kelly Spendlove. A computational framework for the connection matrix theory. *arXiv:1810.04552 [math.AT]*, 2018.

[HT02] Michael A Hunter and Yoshio Takane. Constrained principal component analysis: Various applications. *Journal of Educational and Behavioral Statistics*, 27(2):105–145, 2002.

[Hua20] Mo Huang. *A Statistical Framework for Denoising Single-cell RNA Sequencing Data*. PhD thesis, University of Pennsylvania, 2020.

[HW11] Megumi Harada and Graeme Wilkin. Morse theory for the moment map for representations of quivers. *Geometriae Dedicata*, 150:307–353, 2011.

[JL21] George Jeffreys and Siu-Cheong Lau. Kähler geometry of quiver varieties and machine learning. *arXiv:2101.11487 [math.AG]*, 2021.

[Kac80] V G Kac. Infinite root systems, representations of graphs and invariant theory. *Inventiones Mathematicae*, 56:57–92, 1980.

[Kin94] A D King. Moduli of representations of finite dimensional algebras. *Quarterly Journal of Mathematics*, 45:515–530, 1994.

[Kir16] Alexander Kirillov Jr. *Quiver representations and quiver varieties*, volume 174 of *Graduate Studies in Mathematics*. American Mathematical Society, 2016.

[KP19] Frances Kirwan and Geoffrey Penington. Morse theory without nondegeneracy. *arXiv:1906.10804 [math.DG]*, 2019.

[Lau96] Steffen L Lauritzen. *Graphical models*, volume 17. Clarendon Press, 1996.

[MWH12] Zhaoshi Meng, Ami Wiesel, and Alfred O Hero. Distributed principal component analysis on networks via directed graphical models. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2877–2880. IEEE, 2012.

[Nak01] Hiraku Nakajima. Quiver varieties and tensor products. *Inventiones Mathematicae*, 146:399–449, 2001.

[NBS12] Thanh T Ngo, Mohammed Bellalij, and Yousef Saad. The trace ratio optimization problem. *SIAM review*, 54(3):545–569, 2012.

[Oud15] Steve Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209 of *Mathematical Surveys and Monographs*. The American Mathematical Society, 2015.

[Rao64] C Radhakrishna Rao. The use and interpretation of principal component analysis in applied research. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 329–358, 1964.

[Sch14] R Schiffler. *Quiver Representations*. Number 184 in CMS Books in Mathematics. Springer, 2014.

[Sul18] Seth Sullivant. *Algebraic statistics*, volume 194. American Mathematical Soc., 2018.

[TA18] Tiffany M Tang and Genevera I Allen. Integrated principal components analysis. *arXiv preprint arXiv:1810.00832 [stat.ME]*, 2018.

[TH01] Yoshio Takane and Michael A Hunter. Constrained principal component analysis: a comprehensive theory. *Applicable Algebra in Engineering, Communication and Computing*, 12(5):391–419, 2001.

[Tod18] Yukinobu Toda. Moduli stacks of semistable sheaves and representations of Ext–quivers. *Geometry and Topology*, 22:3083–3144, 2018.

[TS91] Yoshio Takane and Tadashi Shibayama. Principal component analysis with external information on both subjects and variables. *Psychometrika*, 56(1):97–120, 1991.

[VL76] Charles F Van Loan. Generalizing the singular value decomposition. *SIAM Journal on Numerical Analysis*, 13(1):76–83, 1976.

[WH09] Ami Wiesel and Alfred O Hero. Decomposable principal component analysis. *IEEE Transactions on Signal Processing*, 57(11):4369–4377, 2009.