Principal Components along Quiver representations

Nilava Metya

Rutgers University

December 6, 2023

Table of Contents

The Problem Setting for PCA

Given data $D = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m\} \subseteq \mathbb{R}^n$ with $\frac{1}{m}\sum \boldsymbol{y}_i = \mathbf{0}$, find

1. direction of maximum variance.
2. direction of maximum $r-$variances (that is, directions $\boldsymbol{x}_i$ such that total variance along $\mathbb{R}\langle \boldsymbol{x}_1, \cdots, \boldsymbol{x}_r \rangle$ is maximized).

Solution for 1 direction

The data along direction $\boldsymbol{x}$ (such that $||\boldsymbol{x}|| = 1$) is $\langle D, \boldsymbol{x} \rangle = \{\langle \boldsymbol{y}_i, \boldsymbol{x} \rangle\}_{i=1}^m$. Mean of this data is 0. So, variance of this projected data is $\frac{1}{m} \sum \langle \boldsymbol{y}_i, \boldsymbol{x} \rangle^2 = \sum \boldsymbol{x}^\top \left(\frac{1}{m} \boldsymbol{y}_i \boldsymbol{y}_i^\top \right) \boldsymbol{x} = \boldsymbol{x}^\top \underbrace{\sum \left(\frac{1}{m} \boldsymbol{y}_i \boldsymbol{y}_i^\top \right)}_{S} \boldsymbol{x}$. Thus our problem becomes

$$\max \ \boldsymbol{x}^\top S \boldsymbol{x}$$
$$\text{s.t.} \ \begin{cases} \boldsymbol{x}^\top \boldsymbol{x} = 1 \\ \boldsymbol{x} \in \mathbb{R}^n \end{cases}$$

The optimal $\boldsymbol{x}$ is an eigenvector corresponding to the highest eigenvalue of S.

Main theorem to use

Look at the Rayleigh quotient

$$R_M(\boldsymbol{v}) = \frac{\boldsymbol{v}^\top M \boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}}.$$

Theorem (Rayleigh quotient theorem)

If M is Hermitian, $\max R_M(\boldsymbol{v}) = \lambda_{max}$.

Solution for r directions

Note that $S = \frac{1}{m}\sum \boldsymbol{y}_i \boldsymbol{y}_i^\top$ is symmetric. Assume r distinct eigenvalues $\lambda_1 > \cdots > \lambda_r$.
$\boldsymbol{x}_1 \quad \cdots \quad \boldsymbol{x}_r$.

Highest variance is along $\boldsymbol{x}_1$.
Second highest variance obtained by restricting S to $\langle \boldsymbol{x}_1 \rangle^\perp$ — so direction is $\boldsymbol{x}_2$.
And so on. . . .

This is simply the solution to the optimization problem

$$\max \operatorname{tr}\left(X^\top S X\right)$$
$$\text{s.t. } \begin{cases} X^\top X = \mathbf{1}_r \\ X \in M_{n\times r} \end{cases}$$

Principal Component Analysis *along* A Subspace

Say we have a subspace $V \subseteq \mathbb{R}^n$ with an orthogonal complement U and $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k$ is an orthonormal basis of U. So $W = \begin{bmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k \end{bmatrix} : \mathbb{R}^k \to \mathbb{R}^n$ has image U. Then $V = \left\{ \boldsymbol{x} \in \mathbb{R}^n : W^\top \boldsymbol{x} = \boldsymbol{0} \right\}$.

Then the optimization problem *along* V is

$$\max \operatorname{tr}\left(X^\top S X\right) \qquad \text{implicit}$$

$$\text{s.t.} \begin{cases} X^\top X = \mathbf{1}_r \\ W^\top X = \mathbf{0} \\ X \in M_{n\times r} \end{cases}$$

Alternate perspectives

If $F : \mathbb{R}^d \hookrightarrow \mathbb{R}^n$ is an embedding for V, that is, F is full rank and $\Im(F) = V$.

The X in the above optimization problem would then look like $X = FY$ for some $Y \in M_{d\times r}$.
Then the optimization problem *along* V is

$$\max \operatorname{tr}\left(Y^\top F^\top S F Y\right) \quad \text{s.t.} \quad \begin{cases} Y^\top F^\top F Y = \mathbf{1}_r \\ Y \in M_{d\times r} \end{cases}$$

parameterized

Alsn now notice that $B = FF^\top : \mathbb{R}^n \to \mathbb{R}^n$ is a projection onto V. Then X can be replaced with $X = BZ$ for $Z \in M_{n\times r}$.
Then the optimization problem *along* V is

$$\max \operatorname{tr}\left(Z^\top B S B Z\right) \quad \text{s.t.} \quad \begin{cases} Z^\top B^2 Z = \mathbf{1}_r \\ Z \in M_{n\times r} \end{cases}$$

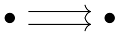
projected

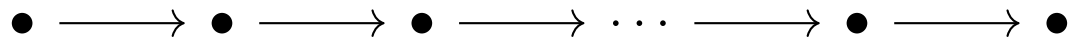
Table of Contents

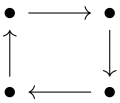
Just graphs...

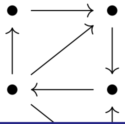
Definition (Quiver)

A quiver is a (finite) directed graph.









Definition (Quiver Representation)

A representation $\boldsymbol{A}_\bullet$ of a quiver $Q = (V, E)$ is an assignment of vector spaces $\boldsymbol{A}_v$ to each vertex $v \in V$ and linear maps $\boldsymbol{A}_e : \boldsymbol{A}_u \to \boldsymbol{A}_v$ to each edge $e : u \to v$.

$$\mathbb{R}^2 \underset{B}{\overset{A}{\rightrightarrows}} \mathbb{R}^3$$

where $A = \begin{bmatrix} 2 & 2 & 2 \\ 1 & 0 & 7 \end{bmatrix}, B = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 1 & 7 \end{bmatrix}$.

Remark: This can be viewed as a representation of an algebra (the path algebra of the quiver Q) — the representation is $\operatorname{Tot}(\boldsymbol{A}_\bullet) = \bigoplus_{v \in V} \boldsymbol{A}_v$.

Want PC along quiver representations

Say we are given data $D = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m\} \subseteq \mathbb{R}^n \underset{\varphi}{\overset{\sim}{\longleftarrow}} \mathrm{Tot}(\boldsymbol{A}_\bullet)$. Where $\boldsymbol{A}_\bullet$ is a representation of Q.

PCA along the *quiver representation* doesn't make sense if it is just normal PCA.

Want: the vector (for principal directions) to respect the representation.
That is, if $\boldsymbol{\gamma} = (\gamma_v)_{v \in V} \in \bigoplus_{v \in V} \boldsymbol{A}_v$ is a direction with $||\varphi(\boldsymbol{\gamma})|| = 1$, then we'd like $\boldsymbol{A}_e(\gamma_x) = \gamma_y$ for all edges $e : x \to y$ in Q.
So we are interested in the **subspace**

$$\Gamma(Q, \boldsymbol{A}_\bullet) = \{(\gamma_v)_{v \in V} \in \mathrm{Tot}(\boldsymbol{A}_\bullet) : \boldsymbol{A}_e(\gamma_x) = \gamma_y \text{ for all edges } e : x \to y\}.$$

Table of Contents

Data Tensor

	40°C			60°C			80°C			100°C			120°C		
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3
$\boldsymbol{x}_1$															
$\boldsymbol{x}_2$															
$\boldsymbol{x}_3$															
$\boldsymbol{x}_4$															
$\boldsymbol{x}_5$															
$\boldsymbol{x}_6$															
$\boldsymbol{x}_7$															
$\boldsymbol{x}_8$															
$\boldsymbol{x}_9$															
$\boldsymbol{x}_{10}$															

$11 \times 3 \times 5$ *contingency table*

p_i are the variables being measured. Each $\boldsymbol{x}_i$ is a sample measurement of p_1, p_2, p_3. Each block is the set of measurements at a different temperature.

If Water...

Parameters different for liquid vs. gas. + different behaviour at boiling point.
Maybe...

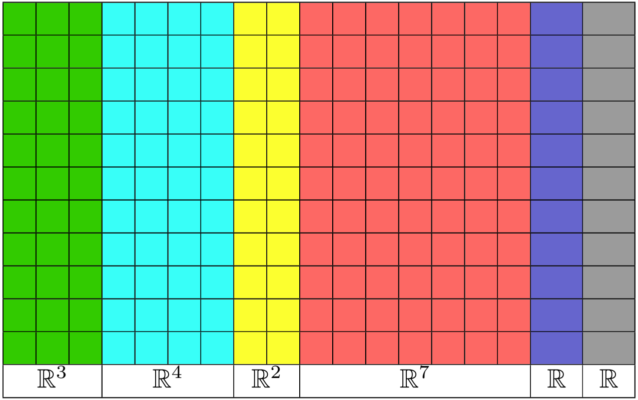
	40°C			60°C			80°C			100°C	120°C	
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	p_2	p_2	p_3
$\boldsymbol{x}_1$												
$\boldsymbol{x}_2$												
$\boldsymbol{x}_3$												
$\boldsymbol{x}_4$												
$\boldsymbol{x}_5$												
$\boldsymbol{x}_6$												
$\boldsymbol{x}_7$												
$\boldsymbol{x}_8$												
$\boldsymbol{x}_9$												
$\boldsymbol{x}_{10}$												

If Water…

Parameters different for liquid vs. gas. + different behaviour at boiling point.
Maybe…

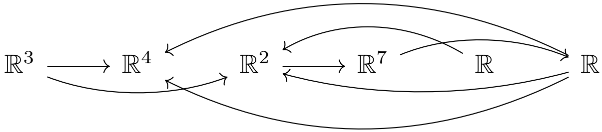
	40°C			60°C			80°C			100°C	120°C	
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	p_2	p_2	p_3
$\boldsymbol{x}_1$												
$\boldsymbol{x}_2$												
$\boldsymbol{x}_3$												
$\boldsymbol{x}_4$												
$\boldsymbol{x}_5$												
$\boldsymbol{x}_6$												
$\boldsymbol{x}_7$												
$\boldsymbol{x}_8$												
$\boldsymbol{x}_9$												
$\boldsymbol{x}_{10}$												
	$\mathbb{R}^3$			$\mathbb{R}^3$			$\mathbb{R}^3$			$\mathbb{R}$	$\mathbb{R}^2$	

In general table could look like...



Why would I group random data? (rhetorical)

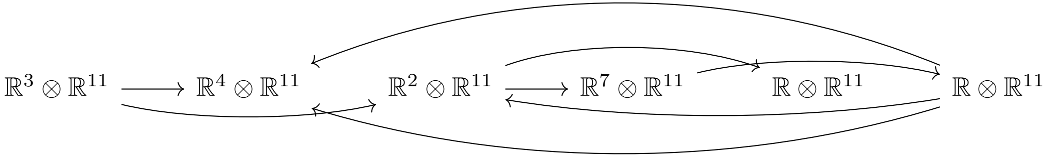
So maybe there's some relation



Restrict to the case when these maps are linear.

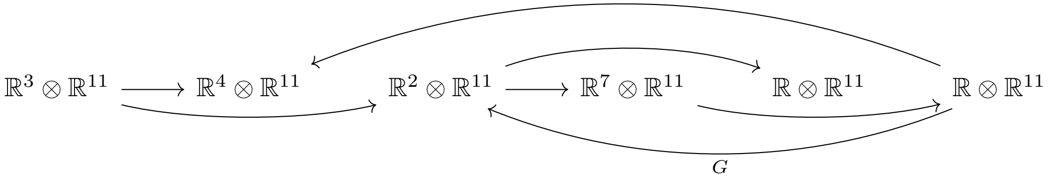
$\mathbb{R}^n$ IS JUST FOR EACH ROW

But the data really look like



Linear $T : U \to W$ induces linear $T \otimes \mathbf{1}_V : U \otimes V \to W \otimes V$.

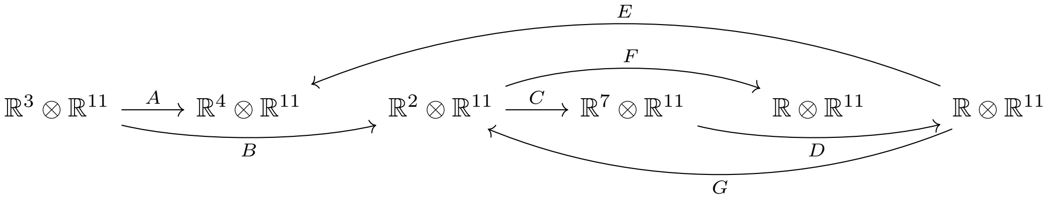
Look at a simpler example



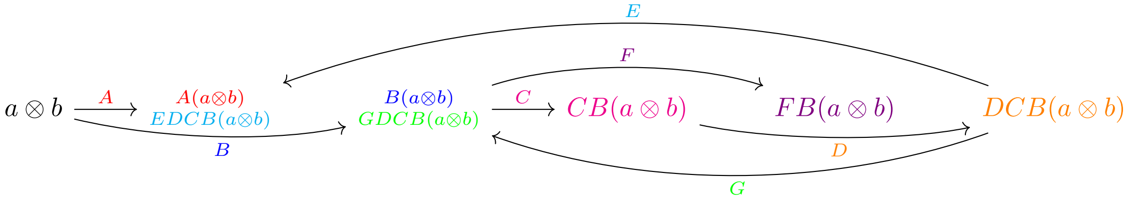
Want a rank 1 approximation of the data in this quiver representation.

Want . . .

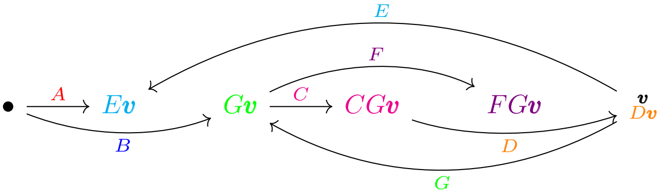
Look at a simpler example



Want a rank 1 approximation of this **quiver representation**. Should respect the representation.



Could we have started elsewhere?



Stuck!!!

Key takeaways/observations

- $\Gamma(Q, \boldsymbol{A}_\bullet)$ is precisely what we want.
- *Strongly connected components* and *minimal vertices* are important.
- Arborescences are good: If Q is an arborescence with root ρ, then $\Gamma(Q, \boldsymbol{A}_\bullet) \cong \boldsymbol{A}_\rho$.
- If there are two paths p, q from u to v, then γ_u is good iff $\gamma_u \in \ker(\boldsymbol{A}_p - \boldsymbol{A}_q)$.

Table of Contents

Ear decomposition

Definition

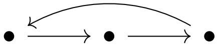
An ear decomposition $Q_\bullet$ of Q is an ordered sequence of $c \geq 1$ subquivers $\{Q_i = (s_i, t_i : E_i \to V_i) | i \in [c]\}$ of Q subject to the following axioms:

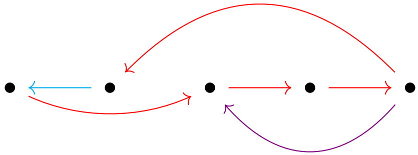
1. the edge sets E_i partition E;
2. the quiver Q_1 is either a single vertex or a cycle, while Q_i for each $i > 1$ is a (possibly cyclic) path in Q;
3. for each $i > 1$, the intersection of V_i with the union $\bigcup_{j<i} V_j$ equals $\{s(Q_i), t(Q_i)\}$.

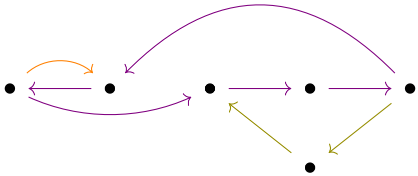
Theorem

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

Examples

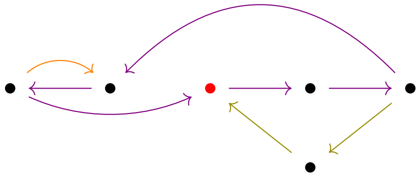




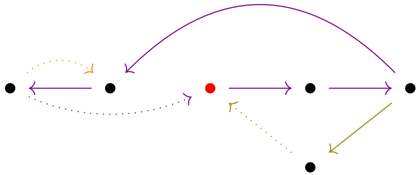


Arborescence from SC quiver (representaiton)

Fix an origin ρ for Q_1.

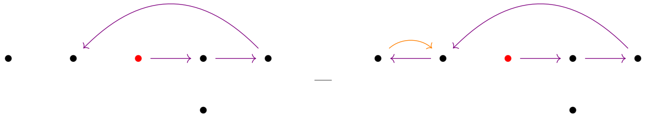


Remove terminal edges to get T.



We get an arboroscence.
For each terminal edge ϵ define $\Delta_\epsilon : A_\rho \to A_{t_\epsilon}$ by $\Delta_\epsilon = A_{p(t_\epsilon)} - A_\epsilon \circ A_{p(s_\epsilon)}$.

For the orange edge:



The root is therefore

$$\left(\boldsymbol{A}|_T\right)_\rho \cong \bigcap_{\epsilon \text{ terminal}} \ker \Delta_\epsilon =: K$$

Lemma

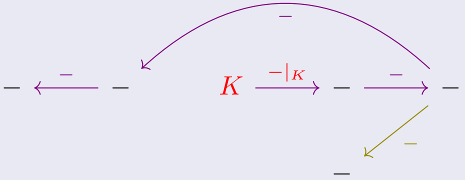
$$\Gamma(Q, \boldsymbol{A}_\bullet) \cong \left(\boldsymbol{A}_\bullet|_T\right)_\rho \cong \bigcap_{\epsilon \; terminal} \ker \Delta_\epsilon$$

What if there's something more than strongly connected?

Want to find a representation of the new quiver (arborescence) whose space of sections is unchanged.

Lemma

This new representation is



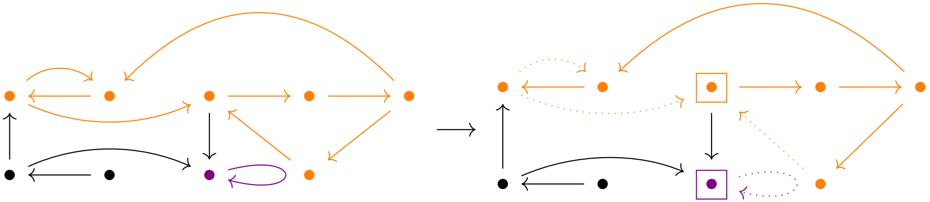
.

Table of Contents

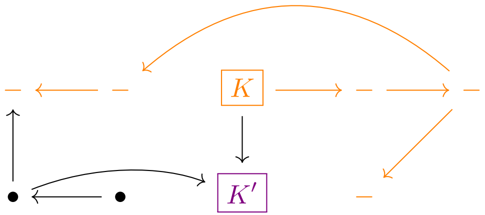
In the quiver

Cycles occur in strongly connected parts. So we

1. Identify each maximally strongly connected component of the quiver, and
2. Get an arborescence from there by removing terminal edges (keep track of the roots).



In the representation



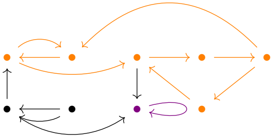
For the $\bullet$ in the $(2,2)$ position, what if after pushing a vector to $(2,3)$ (via $(2,1)$), the vector doesn't fall into K'? So, replace it with $A^{-1}_{p(\bullet \to K')}(K')$. In fact, it needs to be replaced with the intersection of all paths $\bullet \to \rho(R)$ for all strongly connected components R, where $\rho(R)$ is the root of R.
For the $\bullet$ in the $(2,1)$ position, replace it with intersection of all $A^{-1}_{p(\bullet \to \rho(R))}(K(R))$.

Table of Contents

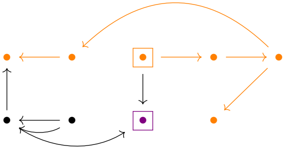
Still not done?

$\bullet \rightrightarrows \bullet$ and $\overset{\curvearrowright}{\bullet \longrightarrow \bullet \longrightarrow \bullet}$ are still problems.

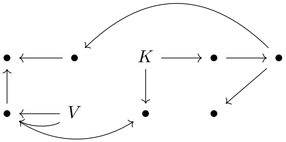
Start with



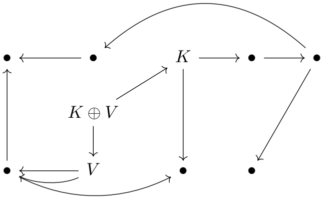
Do an acyclic reduction



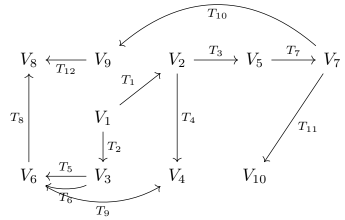
Identify minimal vertices

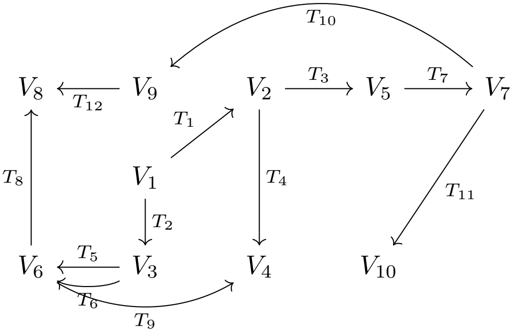


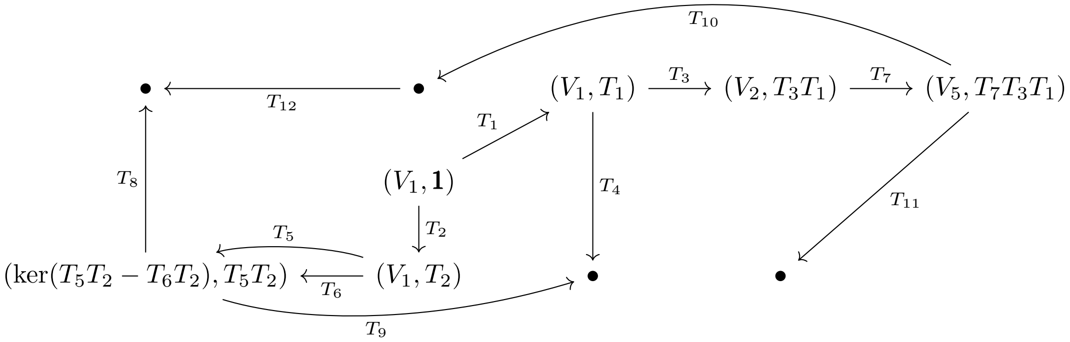
Augment the quiver to have a single source ρ.
Connect a new vertex to minimal vertices.



Now flow through. . .







Replace each $\boldsymbol{A}_v$ with the equalizer Φ_v of $\left\{\boldsymbol{A}_e\varphi_{s(e)} : \Phi'_v \to \boldsymbol{A}_v | e \in E_{\text{in}}(v)\right\}$ with a flow map $\varphi_v : \Phi_v \to A_v$ given by the restriction of $\boldsymbol{A}_e\varphi_{s(e)}$.

Lemma

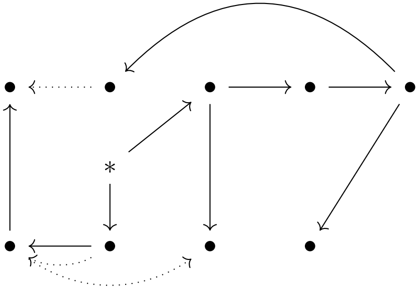
$\Gamma(Q, \boldsymbol{A}_\bullet) \cong \bigcap_{v \in V_{max}} \Phi_v =: Q\Phi(\boldsymbol{A}_\bullet)$.

The resulting quiver representation

Lemma

If $T^+ \subseteq Q^+$ is a spanning arborescence (found by BFS) of the augmented quiver with root ρ, then an arboreal replacement of $\boldsymbol{A}_\bullet$ is the representation given by $\boldsymbol{A}_\bullet^+$ of T^+ given by

$$(\boldsymbol{A}_\bullet^+)_v = \begin{cases} \Phi(A_\bullet) & \text{if } v = \rho \\ \boldsymbol{A}_v & \text{otherwise} \end{cases} \text{ and } (\boldsymbol{A}_\bullet^+)_e = \begin{cases} A_e|_{\Phi(A_\bullet)} & \text{if } s(e) = \rho \\ \boldsymbol{A}_e & \text{otherwise} \end{cases}.$$



Thankyou!