Notes about Network Propagation

Yiftach Kolb

September 6, 2020

1 Mathematical Primer

1.a Matrix Algebra Primer

Definition 1.1. A **Transition Matrix** is a real valued non negative square (n^2) matrix A s.t. each of its columns sums up to one:

$$\forall j \sum_i A_{i,j} = 1,\ \forall i,j A_{i,j} \geq 1$$

A acts from the left as a linear mapping: $T(v) := Av, \forall v \in R^n$. In the following script we might interchangeably and indistinguishably use T (the mapping) or A the matrix.

A transition is **positive**, designated by $A > 0$ if all its entries are positive.

A transition is **regular** if for some $k > A^k$ is positive. The same property is called **primitive** in some other sources.

A transition is **irreducible** if for every entry $A_{i,j}$ there is some k such that $A^k_{i,j} > 0$ (More on that further down).

Equivalently it can be shown that A matrix A is irreducible by if and only if it is NOT similar by permutations matrices to a block upper triangular matrix which means $\nexists P : A = P \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} P^{-1}$

Where P is a permutation matrix and B, C are square matrices of size greater than 0.

Definition 1.2. A **State** is a non-negative vector $v \in R^n$ s.t $\sum_i v_i = 1$.

Remark 1.1. If v is a state and A is a transition as defined in 1.1, then Av is also a state because:

$$\sum_i (Av)_i = \sum_j v_j (\sum_k A_{j,k}) = \sum_j v_j \cdot 1 = 1$$

If A is a transition and x, y are two states such that $x \leq Ay$ then $x = y$.

Definition 1.3. Given $A \in \mathbb{C}^{n \times n}$ We let $|A| \in \mathbb{C}^{n \times n}$ be the resulting matrix from applying $|\cdot|$ element wise. Given a vector $v \in \mathbb{C}^n$ we let $|v| \in \mathbb{C}^n$ the corresponding non-negative vector.

We also let $A > 0, v > 0$ mean that it holds coordinate wise.

Here is a very useful lemma for non-negative matrices which we will need later:

Lemma 1.1. *Let $0 \leq A \leq B \in \mathbb{R}^{n\times n}$ and let $0 < v \in \mathbb{R}^n$. If $Av = Bv$ then $A = B$.*

Proof. Trivial. □

Remark 1.2. If $u \in \mathbb{C}^n$ is on the unit circle and T a transition then $|u|$ is a transition, meaning $\||u|\| = 1$, so $T|u|$ is also a transition so $\|T|u|\| = 1$.

We have (component-wise) $|Tu| \leq T|u|$. If $T > 0$ and u has negative or non-real entries, then this inequality must be strict and then $\|Tu\| < \|T|u|\| = 1$.

Lemma 1.2. *If T is a transition, then there is a state u such that $Tu = u$.*

Proof. Because the columns of A all sum to 1, the columns of $A - I$ all sum to 0. Therefore $(1, 1, \ldots, 1)$ is an (left) eigenvector of the rows with eigenvalue 1. Therefore there is also some real (right) column eigenvector with eigenvalue 1. (Also it follows from the Brouer fixed point theorem because T maps the l_1 sphere to itself).

Let $u \in R^n$ be such vector: $Au = u$. Let $u = u^+ + u^-$ such that $u^- = \min(u_i, 0)$ and the other one defined similarly for the non-negative components.

Because A is non-negative $A(u^+) \geq (Au)^+ \geq 0$, and $A(u^-) \leq (Au)^- \leq 0$.

From A being non-negative and $(Au)^+ + (Au)^- = Au = u = u^+ + u^-$ And also $(Au)^+ = (u)^+$, so we must have $Au^+ \geq (Au)^+ = u^+$ (component wise). But if we had a strict inequality we would get: $\|A(u^+/\|u^+\|_1)\|_1 > 1$ which is a contradiction to A being a transition matrix.

Then $Au^+ = u^+, Au^- = u^-$ and one of them must be non-zero. It follows that A has a non-negative eigenvector with eigenvalue 1 (one of $u^+, -u^-$ which is non-zero). If we l_1-normalize that eigenvector it becomes a state. □

Lemma 1.3. *If a transition $A > 0$ (or primitive) , then it has exactly one eigenvector v with eigenvalue 1 and in addition it can be chosen so that $v > 0$ Any for any other eigenvalue λ it holds that $|\lambda| < 1$.*

If A is just irreducible then then again $v > 0$ and is unique but there may be additional eigenvalues on the unit circle.

Proof. Let $A > 0$ be a transition. We already know that there exists at least one such eigenvector. Let u, v s.t $Au = u, Av = v$. We can assume these are real vectors because A has only real entries. Therefore we can choose u, v to be states as we have already proven above.

Then let $w = u - v$. So $Aw = A(u - v) = u - v = w$. And $\sum_i w_i = 1 - 1 = 0$ by choice of u, v.

Like before $Aw^+ = w^+, Aw^- = w^-$ and because $w \neq 0$ but sums up to 0, both $w^+, -w^- > 0$. Because w^- is non zero exactly on entries where w^+ is zero and vice versa, each of them must have at least one 0 entry (and one none zero). But because A is strictly positive and w^+ is non-negative, Aw^+ must have ONLY positive entries, contradicting $Aw^+ = w^+$. It follows then that $u - v = 0$ is the only possibility, hence the eigenvector with 1 as eigenvalue is unique.

Suppose there is $Aw = \lambda w$ where $|\lambda| = 1$. Choose w so it is l_1-normalized. Then $|w| = |Aw| \leq A \cdot |w|$ If w has any negative or complex coordinantes, then $|Aw| < A|w|$ and therefore $1 = \||Aw|\|_1 < \|A|w|\|_1 = 1$, a contradiction. Therefore there cannot be any other eigenvalues on the unit circle.

Extending this for primitive matrix is easy because for some sufficiently big k $A^k > 0$.

To prove the uniqueness of the 1-eigenvector for the irreducible case, we have $(\forall k \in \mathbb{N}) A^k w^+ = w^+$ and from that with some more work left undone it follows that $w^+ > 0$ or $w^+ = 0$. □

□

Remark 1.3. The lemmas and theorems in this section are phrased in term of transitions. They hold true almost verbatim in the more general case of positive/non-negative linear transformations.

In general a non-negative linear transformation has a **spectral radius** $\rho = \rho(A)$ which is the absolute value of its greatest eigenvalue. In the case of positive maps there is a unique single eigenvector with ρ as the unique greatest eigenvalue and so forth .... When we deal with a transition map, lemma 1.3 guaranties it has a spectral value of $\rho(A) = 1$.

Theorem 1.1. *Let T be a positive (or primitive) transition. Then*

1. 1 is the greatest eigenvalue of T and it has one unique eigenvector which is also positive, so there exists a unique stationary state.

2. All the other eigenvalues have absolute value strictly less than 1.

3. For every state u, $T^k u$ converges to the stationary state π. In particular the columns of T converge to π.

Proof. 1 and 2. We already know.

3. There is a Jordan decomposition $T = PJP^{-1}$.Such that J is a Jordan matrix, $j_{1,1} = 1$ and the rest of the main diagonal $|J_{i,i}| < 1$. So now the convergence is clear $J^k \to (e_1|0\ldots|0)$. For the matrix P it must hold that $P = (P_1, \ldots, P_n)$ and P_1 is the eigenvector of T corresponding to 1 which we are free to l_1 normalize and the first row of P^{-1} is the left eigenvector corresponding to 1 and so force ....

Some more work or literature check should confirm that $T^k \to (v|v\ldots|v) = V$. Then one can verify $Vu = v$ for any state u. □

Theorem 1.2. *Let T be an irreducible positive (or primitive) transition. Then:*

1. Then 1 is the greatest eigenvalue of T and it has one unique eigenvector which is also positive, so there exists a unique stationary state.

2. If there are other other eigenvalues on the unit circle then their algebaic multiplicity is equal their geometric multiplicity.

3. For every state u, $\frac{1}{n}\sum_{k=1}^{n} T^k u$ converges to the stationary state π. In particular the columns of T converge to π.

Proof. 1 We already know.

2. There is a Jordan decomposition $T = PJP^{-1}$. such that J is a Jordan matrix, $j_{1,1} = 1$ and the rest of the main diagonal $|J_{i,i}| \leq 1$.

If we had a Jordan block of size greater than 1 for an eigenvalue λ, Then on the superdiagonal of J^k we would have $k\lambda$. If $\lambda = 1$ then J^k and hence T^k would be unbounded, but that is impossible since T^k is a transition. If follows that all eigenvalues on the unit circle must be simple (alebgaic multiplicity equals geometric).

3. For the convergence, it follows from calculations on the Jordan blocks, which I omit. See Meyer [8] for rigorous proof. □

What differs irreducible non-primitive matrices from being primitiv is that they are periodical on their complex eigenvectors with complex eigenvalues on the unit cycle. There is, in fact a wonderful theorem from Wielandt which characterizes these Matrices:

Theorem 1.3 (Wielandt (1950))**.** *Let $A, B \in \mathbb{C}^{n\times n}$ such that $A \geq 0$ is irreducible and $|B| \leq A$ (component-wise). Then $\rho(B) \leq \rho(A)$. If $\rho(B) = \rho(A)$ then $|B| = A$, B has an eigenvalue of the form $\mu = \rho(A)e^{i\phi}$ and:*

$$B = e^{i\phi} DAD^{-1}$$

where D has the form:

$$D = \begin{pmatrix} e^{\theta_1} & & & \\ & e^{\theta_2} & & \\ & & \ddots & \\ & & & e^{\theta_2} \end{pmatrix} \tag{1}$$

And conversely any B of the form 1 has $\rho(B) = \rho(A)$, $|B| = A$ and μ is an eigenvalue of B which corresponds to the eigenvalue $\rho(A) = |\mu|$ the greatest eigenvalue of A.

Proof. To see a rigorous proof I suggest Meyer [8].

The keys for proving this theorem are: First WLG assume A is a transition. This is possible because we may replace A with $W^{-1}A$, and B with $W^{-1}B$, where W is the diagonal matrix that has the column-sums of A. Since A is irreducible it cannot have a column or a row that is all 0 so this diagonal is positive W is indeed invertible and later we can cancel out the W's and return to the general case.

Let v be the μ-eigenvector $Bv = \mu v, \|\mu| = 1$ and choose it so that $\|v\|_1 = 1$.

Then

$$|v| = |\mu v| = |Bv| \leq |B||v| \leq A|v| \tag{2}$$

Since A is a transition by remark 1.1 $A|v| = |v|$. Since A is irreducible and $A|v| = |v|$ we must have by 1.2 that $|v| > 0$, and then by lemma 1.1 $A = |B|$ so that proves the first part.

Now let $w = v/|v|$ (component-wise division) and let

$$D = \operatorname{diag}(w) = \begin{pmatrix} e^{\theta_1} & & & \\ & e^{\theta_2} & & \\ & & \ddots & \\ & & & e^{\theta_2} \end{pmatrix}$$

Then $v = D|v|$, $|v| = D^{-1}v$ and we have:

$$\begin{aligned} A|v| = |v| = D^{-1}v = \\ = D^{-1}\mu^{-1}Bv = \mu^{-1}D^{-1}BD|v| \end{aligned} \tag{3}$$

We know already that $|v| > 0$. If $C := \mu^{-1}D^{-1}BD$ contains any negative or complex entries, then 3 cannot hold. It follows that

$$A = C = \mu^{-1}D^{-1}BD$$

This proves the harder direction of the claim, the other direction is easy □.

□

The amazing consequence from 1.3:

Theorem 1.4 (Corollarly)**.** *If A is irreducible transition with h eigenvalues on the unit circle then its eigenvalues are exactly the hth unit roots $\lambda_k = e^{2\pi ik/h}, k = 0 \dots n-1$ and A is similar to λA by a diagonal matrix for any such eigenvalue.*

Proof. Use theorem 1.3 with $B = A$. If $|\lambda| = 1$ is an eigenvalue then $A = \lambda DAD^{-1}$. Since similarity doesn't change the eigenvalues, A and λA must have the same eigenvalues with the same multiplicity. Since 1 is a simple eigenvalue of A and hence of λA, and its corresponding eigenvalue λ is simple in λA and therefore in A as well.

The matrices $\lambda_k A$, $k = 0 \dots h$ are all similar, with $\lambda_0 = 1$ and $|\lambda_k| = 1, k = 0\dot{h} - 1$ all simple eigenvalues. The only way for this to hold is if $\{\lambda_k\}_0^{h-1}$ form a multiplicative group of order h on the unit circle, in other words, the eigenvalues are exactly all the hth unit roots. □

Remark 1.4. If A is irreducible transition with exactly h eigenvalues on the unit circle then h is called the **period** of A. $A \geq 0$ is primitive if and only if it is irreducible and aperiodic ($h = 1$).

If $\omega = e^{2\pi i/h}$ then 1.4 shows that $A = \omega DAD^{-1}$. So if λ is any eigenvalue not necessarily on the unit circle, then $\omega\lambda$ is also an eigenvalue with the same multiplicity and rotation with ω is an automorphism on the eigenvalues.

We may choose D so that $D[1,1] = 1$. If we reindex the dimension we can make D look like

$$D = [D_0|\omega D_1| \dots \omega^{h-1}D_{h-1}]$$

So Indexes corresponding to the same phase of the period appear sequentially.

Then use the identify $\forall k D^h A^k = A^k D^h$ and the fact that A is irreducible to show that $D^h = I$. Then use the identity $\omega DA = AD$ to show that in the new indexing A has the following periodical block structure (0 on the main diagonal):

$$A = \begin{pmatrix} 0 & M_1 & 0 & \dots & 0 \\ 0 & 0 & M_2 & 0 \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & 0 & M_{h-1} \\ M_h & 0 & \dots & 0 & 0 \end{pmatrix}$$

Now we will just present the Perron-Frobenius theorems. The main parts that are important to our work have appeared in the previous theorems.

Theorem 1.5 (Perron-Frobenius [8])**.** *Let $0 < A \in \mathbb{R}^{n\times n}$ with spectral radius $\rho := \rho(A)$, then the following are all true:*

- $\rho > 0$
- *ρ is a simple root of the characteristic polynomial of A, in other words its algebraic multiplicity is* 1.
- $(\exists v > 0) Av = \rho v$
- *If $Au = \lambda u$ and $\|u\| = \rho$ then $\lambda = \rho$ namely, ρ is the unique eigenvalue on the spectral circle.*
- *(Collatz-Wielandt Formula) $\rho = \max_{x \in \Gamma} \min_{i:x_i \neq 0} [Ax]_i / x_i$ with $\Gamma = \{x | x \geq 0, x \neq 0\}$*

Theorem 1.6 (Perron-Frobenius for irreducible matrices [8])**.** *Let $0 \leq A \in \mathbb{R}^{n \times n}$ be irreducible with spectral radius $\rho := \rho(A)$, then the following are all true:*

- $\rho > 0$
- *ρ is a simple root of the characteristic polynomial of A, in other words its algebraic multiplicity is* 1.
- $(\exists v > 0) Av = \rho v$
- *There are no additional non-negative unit eigenvectors of A other than v.*
- *(Collatz-Wielandt Formula) $\rho = \max_{x \in \Gamma} \min_{i:x_i \neq 0} [Ax]_i / x_i$ with $\Gamma = \{x | x \geq 0, x \neq 0\}$*

1.b More on Matrices, Graphs and Stochastics

A directed non-weighted graph can be uniquely represented by its adjacency matrix. $A_{i,j} = 1$ if and only if there is a directed edge from i to j (if we want to use it for transitioning on columns as done above). It's possible to assign different edge weights rather than only 1 and 0. If the graph is undirected each edge would be counted in both directions and the matrix is symmetric. Relabeling the vertices of a graph yields an adjacency matrix that is similar by permutations (PAP^{-1}, where P is a permutation matrix) and vice versa.

To turn A into a transition, normalize each column by dividing it with its out going rank, so let $D_{i,i} = \text{outrank}(i)^{-1}$, AD is the transition matrix of this graph (because right-multiplying by D normalizes each column by its rank). If the graph was stochastic to begin with then the adjacency matrix as we defined it is already column-normalized.

Definition 1.4. A graph is G **strongly connected** if there is a directed path form any edge to any edge. if and only if the adjacency matrix is irreducible.

We say that the **Period of a vertex** $v \in V(G)$ is the greatest common divider of all closed paths on v. If the periods of all vertices are equal (spoiler- in the case that G is strongly connected they do), we call it the **Period of the graph** G.

Remark 1.5. If G is strongly connected then the periods of all vertices are indeed equal and its easy to prove. The corresponding adjacency matrix A is irreducible so it too has a period h as defined in 1.4 and it is equal to the graph period (can be shown using 1.4 and 1.4).

So if the graph G is strongly connected and has period 1 then the adjacency matrix is aperiodical and hence primitive, and vice versa.

Primitivity is a stronger condition than irreucibility and also I think it is easy to that if A is irreducible then $I + A$ is regular, because as we exponentiate it every positive entry remains positiv ....

We first show that the two condidtions for irreducibility of a matrix are equivalent. If A is similar to a block upper triangular matrix, than clearly the bottom left 0-block is going to stay all-0s in the similar matrix and the corresponding entries in the original matrix as well. So there are i, j such that $\forall m A^m_{i,j} = 0$

On the other direction if A has indices i, j such that $\forall m A^m_{i,j} = 0$ rename i to 1 and j to n. Now consider the graph G that A is its adjacency matrix. Then rename the vertices so that then first k vertices are all the vertices reachable from 1, and the other are the vertices unreachable from 1 (we know that n is unreachable). Since there is no edge between vertice from $\{1 \dots k\}$ to any vertice of $\{k+1 \dots n\}$ the adjacency matrix is upper block-triangular $\square$.

Now we show that if G is strongly connected then A is irreducible: Let graph G be strongly connected and A ist adjacency matrix. We want to show that it is irreducible. Let G^n be the graph on the same vertices such that i, j have an edge in G^n iff there is a path of length n in G connecting them. Then its adjacency matrix is exactly A^n converted to boolean (positive means true). It is relatively easy to show by induction but I won't show it reigorously. Just consider that if and only if there is a pathe i to k to j then $A_{i,k}A_{k,j} > 0$. So from here we see that if G is strongly connected, then for any i, j we can find some m so that there is is an m-path connecting them and therefore $A^m_{i,j} > 0$.

Finally we show the other direction: if the graph G is not strongly connected we want to show its adjacency matrix is reducible. We take a node i that has a minimal number of reachable nodes. If i is a sink then we switch it's name to 0. We assume by induction that the rest of the nodes have some permutations that results in a triangular matrix and then we trivially extend it with the sink node that we took out as the new first column. If i is not a sink by minimality there must be a cycle of minimally connected nodes including i, so we reduce all of them to one representative which must be a sink of the reduced graph. Now build the block triangular matrix on the smaller graph (so induction hypothesis), then extend it again which should be easy because the removed nodes form a block that connects only to itself (columns with 1 only between said indices, 0 otherwise) so we can put them as the first indices of the matrix ....

It is possible to turn a reducible transition into a regular and therefore irreducible by way of random restart as explained in the article Cowen et al. [1].

If we start with the adjacency matrix and replace each 0 with some small ϵ representing a light weight edge, then normalize the rows, we would get the same thing. Another way to see it: take two transition matrices P and Q, then any convex combination of them, namely $\alpha Q + (1-\alpha)P$ for $\alpha \in [0, 1]$ is also a transition because the rows clearly still all sum to 1.

Remark 1.6. Regarding permutation matrices and similarity. There is a natural isomorphism between the permutation group S_n and the parmutation matrices of $n \times n$ size: $\pi \mapsto (e_{\pi(1)}, \dots, e_{\pi(n)})$. If P is a permutation matrix with a corresponding permutation π, and A any (n square) matrix, then PA is be the matrix obtained by permuting the rows of A according to the permutation π. AP is the result of permuting the columns by the permutation π^{-1}, why is that? see below for the explanation but if you just consider that $P = \prod \Theta_i$ is the product of permutation matrices that each corresponds to a 2-cycle permutations (I think these are called transpositions), it becomes clear.

Now when we are dealing with adjacency matrices, we want to rename the indices, thereby rearranging the adjacency matrix to be block triangular. This 'rearrangement' means exactly multiplying it from left and right by some PAP^{-1}, and from here arises this similarity condition. If we recall every permutation π is a composition of 2-cycles, and 2-cycles are their own inverse. So when we switch indices i and j what we actually do is switching row i with jm then permuting column i with j. Then, we permute indices i' and j', so the row permutation will stack up from the left, the column permutation will stack from the right, and thus arises this PAP^{-1} type of

matrix from the original adjacency matrix.

Lemma 1.4. *If s is a state (column), T is a transition, and $\bar{1}$ is the matrix with all 1's, then $\bar{1}s = (1, 1, \ldots, 1)^t = \mathbf{1}$ and $\bar{1}T = \bar{1}$.*

Proof. Every element of the product matrix is a sum of a column of T, hence 1. s is the same it's like taking just one column of T.

We therefore see that, using the notations of the article [1] Since $\bar{1} \cdot P(t) = \bar{1}$, using $(\frac{(1-\alpha)}{N}\bar{1} + \alpha W)$ as the transition matrix (where W is the transition originated from the normalized adjacency matrix of the original graph), we get:

$$P(t+1) = (\frac{(1-\alpha)}{N}\bar{1} + \alpha W) \cdot P(t) = \frac{(1-\alpha)}{N}\mathbf{1} + \alpha P(t)$$

□

Remark 1.7. Instead of taking the uniform distribution, let s be some state, and let $S = (s, s, \ldots, s)$ a matrix whose columns are all s. Then if p is any state, $Sp = s$. So we can replace $\mathbf{1}/N$ with S in the above remarks, so our random restart distribution can be chosen arbitrarily we actually don't need to require that that it is has no 0 entries:

Let $A, B \neq 0$ be non-negative square matrices, such that A has a positive row, and B is irreducible. Then for all sufficiently large k $(A + B)^k > 0$ and therefore $A + B$ is regular. The proof is pretty straight forward. What's important to our case is that the adjacency matrix of a connected graph has the property of B and the matrix like S has the property of A. Therefore the composite matrix is regular fro any arbitrary restart state s.

Remark 1.8. As we have seen for T regular, and any state s, $T^k s \to \pi$ where π is the unique stationary state.

π can be approximated by itetrating this sequence until sufficent accuracy has been reached.

There is also a direct solution for the random walk with restart which goies as follows: let $p > 0$ be a state (no 0 entries), let W be its column-normalized adjacency matrix, and let $P = (p, \ldots, p)$. as we have seen above, $T = (1 - \alpha)P + \alpha W$ is an irreducible transition matrix therefore π exists and unique.

Let $p_0 = p, p_k := Tp_{k-1} = p_0(1 - \alpha) + \alpha p_{k-1}$ Then $p_k \to \pi$ and We have the relation:

$$I\pi = \lim(p_k) = (1 - \alpha)p_0 + \alpha \lim(Wp_{k-1}) = (1 - \alpha)Ip_0 + \alpha W\pi$$

and rearragement gives:

$$(I - \alpha W)\pi = (1 - \alpha)p_0$$

Now because W is a transition and $0 < \alpha < 1$, for any v such that $\|v\|_1 = 1$, then $\tilde{v}$ is a state and it holds that:

$$\|\alpha W v\|_1 \leq \alpha \|W\tilde{v}\| = \alpha < 1$$

This guaranties that $I - \alpha W$ is invertible and the direct solution is:

$$\pi = (1 - \alpha)(I - \alpha W)^{-1} p_0$$

2 Toying around with graph propagation

2.a Motivation

We work with random graphs genertated by one of the standard models (Erdos- ...) because the real world graphs have similar topology.

For a given category we expect only a small number of proteins play a direct role and hence should carry its annotation, some of them are already known. Ideally we want that with some choice of α (the restart factor), We get a stationary distribution that is fairly concentrated exactly on the proteins that belong to that category. Is it possible?

Even stronger evidence: we start with 2 or more mutally exclusive categories and partial annonations for each, and we want that the stationary distributions will reflect that- namely they should be concentrated on non-overlapping subsets of graph nodes.

3 The Markov Random Fields Methd

We use almost the same terminology of Deng et al. [2] so I won't repeat everything here.

We have a PPI and partial annotaion which gives the potential function $U(x)$ which actually depends on the parameter $\Theta = (\alpha, \beta, \gamma)$ And we want to maximize:

$$P(X|\Theta) = \frac{1}{Z(\Theta)} \exp(-U(x)) \tag{4}$$

To eliminate $Z(\Theta)$ from the equation we are looking at:

$$P(X_i = 1|X_{[-i]}, \Theta) = \frac{\exp(\alpha + (\beta - 1)M_0^i + (\gamma - \beta)M_1^i)}{1 + \exp(\alpha + (\beta - 1)M_0^i + (\gamma - \beta)M_1^i)} \tag{5}$$

If we think of every node labeled with 1 as being occupied by a traveling agent, who in the next iteration is going to visit some other (i.e not allowed to stay in the same place) node according to some probability, then $P(X_i = 1|X_{[-i]}, \Theta)$ is the probability That node i will be visited in the next iteration. In the propagation model we look at the neighbours of i that have a visitor in them, and we look for each of these how many other neighbours it has to calculate the visit probability of i, whereas in equation (5) we don't explicitly look at neighbours of neighbours but rather consider direct interaction between i and all its neighbours.

In the propagation model without restart, we have (of course with the resulting value being capped at a maximum of 1):

$$P(X_i = 1|X_{[-i]}) = \sum_{j \in \text{Nei}(i)} \frac{1_{X_J}}{\text{Nei}(j)} \tag{6}$$

Here we assume that the visiting agent will choose one of the neighbours at equal probability. We can also formulate this similarly to the random restart case. Also perhaps it would make sense, to consider instead of $X_{[-i]}$, a label assignment of all the nodes including i, then update X_i in the next iteration by the probability that it would be visited in the next iteration. And in this case we can add restart an reformulate (6) as:

$$P(X_i = 1|X) = (1-\alpha) \sum_{j \in \text{Nei}(i)} \frac{1_{X_J}}{\text{Nei}(j)} + \alpha \sum_{j:X_j=1} \frac{1}{\text{Nei}(j)} \tag{7}$$

$P(X_i = 1|X)$ here means that giving a labeling X to all the nodes, we calculate the probabiliyt that X_i will be labeled 1 in the next iteration which is the probabilty that at least one of the agents will choose to visit it.

3..1 Estimation of the parametes

The authors of Deng et al. [2] used the quasi-likelihood approach and logistic regression.

If I understand this correctly, they take the subnetwork of annotated proteins (of a specific function), that gives a binary vector $X = (X_i)_{i=1...m}$ and each X_i is treated as an observation that is independent from the other observations, then they find the paramers that best fit the logistic model to the sample distribution.

3..2 Estimating the probabvilty

In the article, for a given a function annotaion, π is the probability that a protein has that a protein has that function annotation (disregarding the information from the PPI network). This π is used to assign random values to the missing data.

An Idea: Maybe we can use use the pagerank (and a cutoff) for the initial assignment instead?

□

4 testing

4.a citations

For more info see (textcite) Meyer [8, p. 631].

This is a normal (cite) [5, p. 115].

This is an (autocite) see [7, p. 231].

Change autocite in the usepackage definition to suit your needs. Use textcite if you want to insert inline cite that includes the author's name.

Here is a [10] cite from wikipedia.

Just cite is just the normal citation command and nocite will make sure that the reference appears in the bibliography even though it may have not specifically been mentioned in the text somewhere.

4.b Images

Here is an image

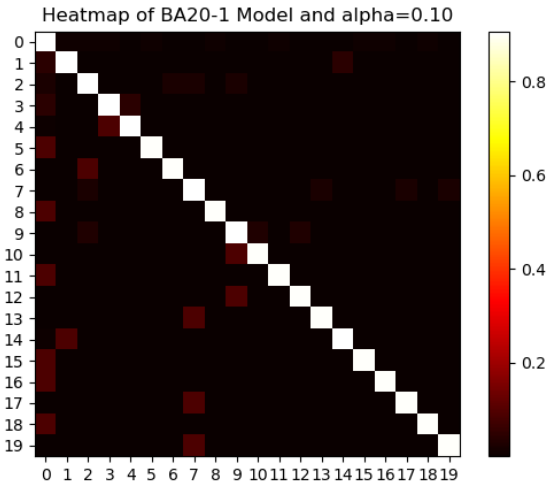

Figure 1: Heatmap. Alpha=0.1

Refer to it with fig 1 tada.

4.c Tables

Table 1: Nernstsche Gleichgewichtspotenziale und das resultierende Membranpotenzial nach Goldmann Gleichung

Ion	rel' Permeabilität	Konz. in	Konz. auß	GG Potenzial
K^+	1	124	4	-86.74
Na^+	0.04	50	470	56.6
Cl^-	0.3	55	580	-59.51
V_m	-51.35			

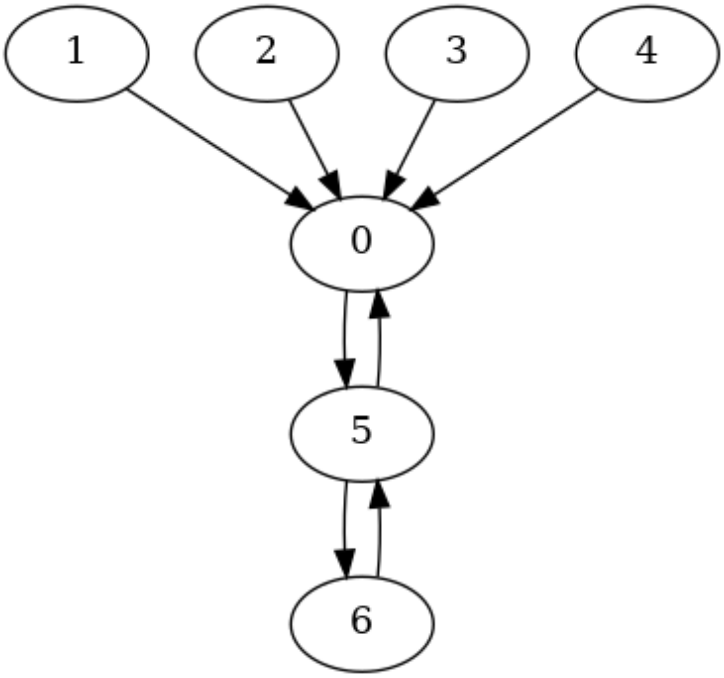

Figure 2: The apparatus used in the experiment. The forward right wing is fixed to the chain wheel which controls its elevation. The electrodes sticking out of the black precision clamp are positioned to touch the exposed meso N1 at the root of the wing

5 Reference

References

[1] Lenore Cowen et al. "Network propagation: a universal amplifier of genetic associations". In: *Nature Reviews Genetics* 18.9 (2017), p. 551.

[2] Minghua Deng et al. "Prediction of protein function using protein-protein interaction data". In: *Proceedings. IEEE Computer Society Bioinformatics Conference*. IEEE. 2002, pp. 197–206.

[3] Michelle Girvan and Mark EJ Newman. "Community structure in social and biological networks". In: *Proceedings of the national academy of sciences* 99.12 (2002), pp. 7821–7826.

[4] Israel Nathan Herstein and David J Winter. *Matrix theory and linear algebra*. Macmillan Publishing Company, 1989.

[5] Hugh Jass. "A big paper". In: *The journal of big papers* MCMXCVII (7991).

[6] Gregory F Lawler. *Random walk and the heat equation*. Vol. 55. American Mathematical Soc., 2010.

[7] Laura Lion, Gabrielle Giraffe, and Carl Capybara. *The dangers of asking the wrong question*. publishing house, 2010.

[8] Carl D Meyer. *Matrix analysis and applied linear algebra*. Vol. 71. Siam, 2000.

[9] Mark EJ Newman. "Modularity and community structure in networks". In: *Proceedings of the national academy of sciences* 103.23 (2006), pp. 8577–8582.

[10] *Perron-Frobenius Theorem*. Wikipedia. 2020. URL: `https://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem` (visited on 08/30/2020).

[11] Martin Vingron. “Propagation methods (Lecture Slides)”.