Matrix multiplication and path counting

Thibaud Taillefumier

The free-living worm *Caenorhabditis elegans* has been proposed by Sydney Brenner as a model organism in part because each worm has a well-defined number of cells (959 in the adult hermaphrodite; 1031 in the adult male). The developmental fate of each of these cells has been mapped and the neuronal connectivity of *C. elegans* (302 neurons) has been elucidated via electronic microscopy (see http://www.wormatlas.org/neuronalwiring.html). Although still a matter of debate, it is worth noting that *C. elegans* neurons appear not to generate action potentials. Moreover, the electronic microscopy data does not unambiguously specify whether synapses are excitatory or inhibitory. Independent of these points, it turns out that matrix operations play a crucial role in exploring biological network data similar to that of *C. elegans* connectome. The aim of this note is to establish the connection between network structure and matrix multiplication.

Connectivity data like that of *C. elegans*, i.e. the exhaustive list of synapses established between anatomically distinguished neurons, can be summarized by an adjacency matrix. Specifically, suppose we labelled N neurons by an index i running from 1 to N. The adjacency matrix A associated with the connectome is defined component-wise as follows: $A_{ij} = 1$ if neuron i connects to neuron j and $A_{ij} = 0$ otherwise. For a network of 5 mapped neurons, we may obtain a matrix A of the form:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} .$$

The above adjacency matrix A indicates that when considered as a pre-synatpic neuron, neuron 1 connects to neuron 2 and neuron 5, and when considered as a post-synaptic neuron, neuron 1 is connected by neuron 3. Note that the row index i is the neuron index when considered as pre-synaptic neuron, whereas the column index j is the neuron index when considered as a post-synaptic neuron. Note also that diagonal entries of A are zero because we consider neurons that do not selfsynapse. Random adjacency matrix can be easily generated via Matlab using the following commands:

```
% create a binary square 5x5 matrix whose entry are 1
with probability 0.3
A=double(rand(5)<0.7);

% substract the diagonal elements
A=A-diag(diag(A))
```

In graph theory, adjacency matrices such as A recapitulate the topology of a directed graph: a directed graph is defined by a list of nodes (e.g. neurons i and j) together with a collection of directed edges between nodes (e.g. synapse $i \to j$).

Let us now interpret matrix multiplication, or rather matrix exponentiation, with respect to the graph structure of the connectome. If the adjacency matrix A is a $N \times N$ matrix, the square adjacency matrix A^2 is a $N \times N$ matrix whose components satisfy:

$$A^2_{ij} = \sum_{k=1}^{N} a_{ik} a_{kj} \, .$$

As usual, these coefficients are the sum of products. What is specific to adjacency matrices is that the intervening coefficients a_{ij} are binary, i.e. either zero or one. This means that $a_{ik} a_{kj} \neq 0$ if and only if $a_{ik} = 1$ and $a_{kj} = 1$, that is if there is a 2-step synaptic route $i \to k \to j$. In other words, the coefficient A^2_{ij} counts the number of 2-step synaptic routes from i to j. From there, we can convince ourselves that this path interpretation generalizes to higher power of the adjacency matrix. For instance, the coefficient

$$A^3_{ij} = \sum_{l=1}^{N} A^2_{il} a_{lj} = \sum_{k=1}^{N} \sum_{l=1}^{N} a_{ik} a_{kl} a_{lj} \, .$$

counts the number of 3-step synaptic routes from i to j. This is already an interesting take on matrix exponentiation. Let us push the argument further by asking whether we can answer the following question: is there a synaptic route from neuron i to neuron j? Stated otherwise, can neuron i influence neuron j? Perhaps surprisingly, this question can be answered practically by introducing the notion of the exponential of a matrix. For real number, the exponential function $e^x = \exp(x)$ is equal to its Taylor series and can be represented as an infinite sum of powers of x:

$$\exp(x) = \sum_{i=1}^{\infty} \frac{x^n}{n!} \, .$$

We can generalize the exponential function to the square matrices by defining the exponential of the $N \times N$ matrix A as the $N \times N$ matrix

$$\exp(A) = \sum_{i=1}^{\infty} \frac{A^n}{n!} \,,$$

where A^n is n-th power of A. Now that we have defined the exponential of the adjacency matrix, we can state the alluded result: there is a synaptic route from neuron i to j, $i \neq j$, if and only if the coefficient $\exp(A)_{ij} \neq 0$. To see why this statement is true, let us suppose that there is a path between i and j and let us denote its length as n. Necessarily, A^n_{ij}, the ij-th coefficient of the n-th power of A, is at least equal to one. Since the coefficients of powers of adjacency matrices are non-negative integers, the definition of $\exp(A)$ tells us that $\exp(A)_{ij}$ is at least equal to $1/n!$, and thus $\exp(A)_{ij} > 0$. Conversely, if there is no path between i and j, the coefficient A^n_{ij} is zero for all n, and the definition of $\exp(A)$ tells us that $\exp(A)_{ij} = 0$. Matlab has a routine that efficiently computes exponential of matrices called `expm`. For the matrix A that we have previously considered, Matlab returns:

```
expm(A)= 1.0000    1.3683         0    1.1782    1.3683
              0    1.5891         0    1.3683    0.5891
         1.0000    1.9574    1.0000    1.5465    1.9574
              0    1.3683         0    2.1782    1.3683
              0    0.5891         0    1.3683    1.5891
```

Such a result indicates that:

Neuron 1 influences neurons 2, 4 and 5 but neuron 1 is only influenced by neuron 3. Neuron 2 influences neurons 4 and 5 and neuron 2 is influenced by all other neurons. Neuron 3 influences all neurons but neuron 3 is influenced by no other neurons. Neuron 4 influences neurons 2 and 5 and neuron 4 is influenced by all other neurons. Neuron 5 influences neurons 2 and 4 and neuron 5 is influenced by all other neurons. Notice that we have disregarded diagonal terms in our discussion of influences.