FUNCTIONS OF SQUARE MATRICES

are defined by expanding the function in the usual (Taylor) manner:

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + ...$$

and replacing x with a square matrix $\mathbb{A}$. Realizing that a square matrix, substituted into its own **characteristic polynomial**, yields $\mathbb{O}$ (zero matrix), its obvious that we should not need more than the first n (where n is the size of the matrix) powers of $\mathbb{A}$ (i.e. $\mathbb{A}^0 = \mathbb{I}$, $\mathbb{A}$, $\mathbb{A}^2$, ..., $\mathbb{A}^{n-1}$). The remaining powers can be obtained, recursively, based on these.

The way how it all works out is as follows: To compute a function of $\mathbb{A}$, we need to find the n eigenvalues of $\mathbb{A}$, say ω_1, ω_2, ..., ω_n (some of them my be complex, some may be multiple) and, for each of these, compute the corresponding **constituent matrix** (we will call these $\mathbb{C}_1$, $\mathbb{C}_2$, ...). Assuming that all eigenvalues are distinct, we can then evaluate any function of $\mathbb{A}$ by

$$f(\mathbb{A}) = \sum_{i=1}^{n} f(\omega_i)\mathbb{C}_i$$

The easiest way to find the constituent matrices is to make the above formula true for $f(x) = x^i$, $i = 0, 1, 2, ..., n-1$, thus:

$$\begin{aligned} \mathbb{C}_1 + \mathbb{C}_2 + \mathbb{C}_3 + ... + \mathbb{C}_n &= \mathbb{I} \\ \omega_1\mathbb{C}_1 + \omega_2\mathbb{C}_2 + \omega_3\mathbb{C}_3 + ... + \omega_n\mathbb{C}_n &= \mathbb{A} \\ \omega_1^2\mathbb{C}_1 + \omega_2^2\mathbb{C}_2 + \omega_3^2\mathbb{C}_3 + ... + \omega_n^2\mathbb{C}_n &= \mathbb{A}^2 \\ .... &= ... \\ \omega_1^{n-1}\mathbb{C}_1 + \omega_2^{n-1}\mathbb{C}_2 + \omega_3^{n-1}\mathbb{C}_3 + ... + \omega_n^{n-1}\mathbb{C}_n &= \mathbb{A}^{n-1} \end{aligned}$$

solving these as if the unknowns and the RHS elements were ordinary numbers.

Example:

$$\mathbb{A} = \begin{bmatrix} 2 & 4 \\ 1 & -1 \end{bmatrix}$$

Characteristic polynomial is $\omega^2 - \omega - 6$, the eigenvalues are $\omega_1 = 3$ and $\omega_2 = -2$. Thus

$$\begin{aligned} \mathbb{C}_1 + \mathbb{C}_2 &= \mathbb{I} \\ 3\mathbb{C}_1 - 2\mathbb{C}_2 &= \mathbb{A} \end{aligned}$$

Solution:

$$\begin{aligned} \mathbb{C}_1 &= \frac{\mathbb{A} + 2\mathbb{I}}{5} = \begin{bmatrix} \frac{4}{5} & \frac{4}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} \\ \mathbb{C}_2 &= \frac{3\mathbb{I} - \mathbb{A}}{5} = \begin{bmatrix} \frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & \frac{4}{5} \end{bmatrix} \end{aligned}$$

We can now evaluate any function of $\mathbb{A}$, e.g.

$$e^{\mathbb{A}} = \begin{bmatrix} \frac{4}{5} & \frac{4}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} e^3 + \begin{bmatrix} \frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & \frac{4}{5} \end{bmatrix} e^{-2} = \begin{bmatrix} 16.095 & 15.96 \\ 3.99 & 4.125 \end{bmatrix}$$

and

$$\mathbb{A}^{-1} = \begin{bmatrix} \frac{4}{5} & \frac{4}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} /3 - \begin{bmatrix} \frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & \frac{4}{5} \end{bmatrix} /2 = \begin{bmatrix} \frac{1}{6} & \frac{2}{3} \\ \frac{1}{6} & -\frac{1}{3} \end{bmatrix}$$

(check).

Multiple eigenvalues

result in a slightly more complicated formula. Suppose that ω_1, ω_2 and ω_3 have the same value. Then, instead of

$$f(\mathbb{A}) = f(\omega_1)\mathbb{C}_1 + f(\omega_2)\mathbb{C}_2 + f(\omega_3)\mathbb{C}_3 + ...$$

we have to use

$$f(\mathbb{A}) = f(\omega_1)\mathbb{C}_1 + f'(\omega_1)\mathbb{D}_2 + f''(\omega_1)\mathbb{E}_3 + ...$$

and, to find the constituent matrices, we now use:

$$\begin{aligned}
&\mathbb{C}_1 + \mathbb{C}_4 + ... + \mathbb{C}_n = \mathbb{I} \\
&\omega_1\mathbb{C}_1 + \mathbb{D}_2 + \omega_4\mathbb{C}_4 + ... + \omega_n\mathbb{C}_n = \mathbb{A} \\
&\omega_1^2\mathbb{C}_1 + 2\omega_2\mathbb{D}_2 + 2\mathbb{E}_3 + \omega_4^2\mathbb{C}_4 + ... + \omega_n^2\mathbb{C}_n = \mathbb{A}^2 \\
&.................. \\
&\omega_1^{n-1}\mathbb{C}_1 + (n-1)\omega_2^{n-2}\mathbb{D}_2 + (n-1)(n-2)\omega_3^{n-3}\mathbb{E}_3 \\
&\qquad\qquad + \omega_4^{n-1}\mathbb{C}_4 + ... + \omega_n^{n-1}\mathbb{C}_n = \mathbb{A}^{n-1}
\end{aligned}$$

which are solved in exactly the same manner as before.

Example:

$$\mathbb{A} = \begin{bmatrix} 4 & 7 & 2 \\ -2 & -2 & 0 \\ 1 & 5 & 4 \end{bmatrix}$$

and has a triple eigenvalue of 2. We get

$$\begin{aligned}
&\mathbb{C} = \mathbb{I} \\
&2\mathbb{C} + \mathbb{D} = \mathbb{A} \\
&4\mathbb{C} + 4\mathbb{D} + 2\mathbb{E} = \mathbb{A}^2
\end{aligned}$$

which yields $\mathbb{C} = \mathbb{I}$,

$$\mathbb{D} = \mathbb{A} - 2\mathbb{I} = \begin{bmatrix} 2 & 7 & 2 \\ -2 & -4 & 0 \\ 1 & 5 & 2 \end{bmatrix}$$

and

$$\mathbb{E} = \frac{1}{2}\mathbb{A}^2 - 2\mathbb{D} - 2\mathbb{I} = \frac{1}{2}\begin{bmatrix} 4 & 7 & 2 \\ -2 & -2 & 0 \\ 1 & 5 & 4 \end{bmatrix}^2 - 2\begin{bmatrix} 3 & 7 & 2 \\ -2 & -3 & 0 \\ 1 & 5 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} -4 & -2 & 4 \\ 2 & 1 & -2 \\ -3 & -\frac{3}{2} & 3 \end{bmatrix}$$

Thus,

$$\mathbb{A}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 2 & 7 & 2 \\ -2 & -4 & 0 \\ 1 & 5 & 2 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} -4 & -2 & 4 \\ 2 & 1 & -2 \\ -3 & -\frac{3}{2} & 3 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -\frac{9}{4} & \frac{1}{2} \\ 1 & \frac{7}{4} & -\frac{1}{2} \\ -1 & -\frac{13}{8} & \frac{3}{4} \end{bmatrix}$$

(check).

TCMC without and with absorbing state(s)

For these, the only function of $\mathbb{A}$ (which is, in this case, the infinitesimal generator of the process with rather special properties, e.g. it must have at least one 0 eigenvalue) is

$$\exp(t\mathbb{A})$$

where t is time. Note that $t\mathbb{A}$ has the same eigenvalues (multiplied by t) and constituent matrices as $\mathbb{A}$ itself, the only thing which changes is thus the e^{ω_j} part (which becomes $e^{t\omega_j}$ when dealing with $t\mathbb{A}$).

Actually, Maple makes it even easier for us; there is a special '`exponential(A,t)`' command (in the 'linalg' package) which evaluates $\exp(t\mathbb{A})$ for us, without having to mess up with any eignevalues, etc. Using this feature of Maple, dealing the a Time-Continuous Markov Chain with any number of states becomes quite trivial (in elements of $\exp(t\mathbb{A})$, and their $t \to \infty$ limit, we can alway find all the answers we need).

When TCMC has one or more absorbing states, the same $\exp(t\mathbb{A})$ enables us to find probabilies of absorption (before time t, or 'ultimately'), and the distribution of the time till absorption (to do this, we would have to pool all absorbing states into one).

`Example:`

$$\mathbb{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & -8 & 2 & 3 \\ 1 & 4 & -6 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

meaning that the first and last states are absorbing. The eigenvalues are 0, 0, -4 and -10. We thus know that

$$f(\mathbb{A}) = f(0)\mathbb{C}_1 + f'(0)\mathbb{D}_1 + f(-4)\mathbb{C}_3 + f(-10)\mathbb{C}_4$$

Using $f(x) = 1$, x, x^2 and x^3 yields

$$\begin{aligned}
\mathbb{C}_1 + \mathbb{C}_3 + \mathbb{C}_4 &= \mathbb{I} \\
\mathbb{D}_1 - 4\mathbb{C}_3 - 10\mathbb{C}_4 &= \mathbb{A} \\
16\mathbb{C}_3 + 100\mathbb{C}_4 &= \mathbb{A}^2 \\
-64\mathbb{C}_3 - 1000\mathbb{C}_4 &= \mathbb{A}^3
\end{aligned}$$

respectively. Solving (ignoring the fact that the unknowns are matrices) results in

$$\begin{aligned}
\mathbb{C}_1 &= \mathbb{I} - \tfrac{39}{400}\mathbb{A}^2 - \tfrac{7}{800}\mathbb{A}^3 \\
\mathbb{D}_1 &= \mathbb{A} + \tfrac{7}{20}\mathbb{A}^2 + \tfrac{1}{40}\mathbb{A}^3 \\
\mathbb{C}_3 &= \tfrac{5}{48}\mathbb{A}^2 + \tfrac{1}{96}\mathbb{A}^3 \\
\mathbb{C}_4 &= -\tfrac{1}{150}\mathbb{A}^2 - \tfrac{2}{600}\mathbb{A}^3
\end{aligned}$$

Substituting powers of $\mathbb{A}$, we get

$$\begin{aligned}
\mathbb{C}_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
\mathbb{D}_1 &= \mathbb{O} \\
\mathbb{C}_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} & -\frac{2}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbb{C}_4 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{6} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

$\exp(\mathbb{A}t)$ is thus computed from

$$\mathbb{C}_1 + \mathbb{C}_3 e^{-4t} + \mathbb{C}_4 e^{-10t}$$

The elements of the first and last column of $\mathbb{C}_1$ (the limit of the previous expression when $t \to \infty$) yield the probabilities of ultimate absorption in first (last) state, given the initial state.

The sum of the first and last column of $\exp(\mathbb{A}t)$, namely

$$\begin{bmatrix} 1 \\ 1 - \frac{2}{3}e^{-4t} - \frac{1}{3}e^{-10t} \\ 1 - \frac{4}{3}e^{-4t} + \frac{1}{3}e^{-10t} \\ 1 \end{bmatrix}$$

provide the distribution function of time till absorption (regardless whether it is into first or last state), given the initial state. Based on that, we can answer any probability question, find the mean and standard deviation of time till absorption, etc.

`Example:` Given we start in the second state, what is the probability that absorption will take more than 0.13 units of time. Answer: $\frac{2}{3}e^{-0.52} + \frac{1}{3}e^{-1.3} = 48.72\%$.

Given we start in the third state, what is the expected time till absorption (regardless where) and the corresponding standard deviation. Answer:

$$\begin{aligned}
\int_0^\infty t\left(\tfrac{16}{3}e^{-4t} - \tfrac{10}{3}e^{-10t}\right)dt &= 0.3 \\
\sqrt{\int_0^\infty (t-0.3)^2\left(\tfrac{16}{3}e^{-4t} - \tfrac{10}{3}e^{-10t}\right)dt} &= 0.2646
\end{aligned}$$