1 Numerical solution of equilibrium equations: direct methods

In this chapter we focus on the numerical solution of the equilibrium equations of a Markov chain with finitely many states.

Consider an irreducible Markov chain with finite state space $\{0, 1, \ldots, N\}$ and transition probability matrix P with elements p_{ij}, $i, j = 0, 1, \ldots, N$. We are interested in calculating the equilibrium distribution $\{p_0, p_1, \ldots, p_N\}$, which is the unique solution to the system

$$p_i = \sum_{j=0}^{N} p_j p_{ji}, \qquad i = 0, 1, \ldots, N, \tag{1}$$

$$\sum_{i=0}^{N} p_i = 1, \tag{2}$$

or in vector-matrix notation, where p denotes the row vector of equilibrium probabilities and e the column vector with ones,

$$p = pP, \qquad pe = 1. \tag{3}$$

The equations (1) are called the equilibrium (or balance) equations. There are several approaches to numerically solve the equilibrium equations. We distinguish between *direct* (exact) methods and *iterative* (approximative) methods. In the next sections we discuss direct methods; the following chapter treats iterative methods.

1.1 Gaussian elimination

The equilibrium equations (1) may be rewritten in the form

$$\sum_{j=0}^{N} p_j (p_{ji} - \delta_{ji}) = 0, \qquad i = 0, 1, \ldots, N, \tag{4}$$

where δ_{ij} is 1 if $i = j$ and 0 else. To find the probabilities p_i, we may apply Gaussian elimination (cf. [2, 3]): we solve p_0 from equation 0 for p_0, and eliminate p_0 from all other equations. Then we solve equation 1 for p_1 and eliminate p_1 from all other equations, and so on, until we solve equation $N - 1$ for p_{N-1}. The last equation is redundant and thus can be omitted (Why?). Hence, if we denote by $a_{ij}^{(n)}$ the coefficients in the equations after n elimination steps, then

$$a_{ij}^{(0)} = p_{ji} - \delta_{ji},$$

and for $n = 0, 1, \ldots, N - 1$,

$$\begin{aligned} a_{nj}^{(n+1)} &= -\frac{a_{nj}^{(n)}}{a_{nn}^{(n)}}, \qquad j > n, && (5) \\ a_{ij}^{(n+1)} &= a_{ij}^{(n)} + a_{in}^{(n)} a_{nj}^{(n+1)}, \qquad i, j > n && (6) \end{aligned}$$

Relation (5) corresponds to the solution of p_n from equation n and relation (6) corresponds to the elimination of p_n from all other equations. Then we calculate

$$v_N = 1, \qquad v_n = \sum_{j=n+1}^{N} a_{nj}^{(n+1)} v_j, \quad n = N-1, N-2, \ldots, 0.$$

The v_n's can be interpreted as *relative visiting frequencies*, i.e., the mean number of the visits to state n between two successive visits to state N. Normalisation of the v_n's finally yields the equilibrium probabilities, so

$$p_i = \frac{v_i}{\sum_{j=0}^{N} v_j}, \qquad i = 0, 1, \ldots, N.$$

The algorithm is summarized in Figure 1; note that we applied Gaussian elimination without pivoting.

Algorithm I (Gaussian elimination without pivoting)

0. Set $a_{ij} = p_{ji} - \delta_{ji}$ for all i and j.

1. Compute for $n = 0, 1, \ldots, N-1$

$$\begin{aligned} a_{nj} &= -\frac{a_{nj}}{a_{nn}}, \qquad j > n, \\ a_{ij} &= a_{ij} + a_{in} a_{nj}, \qquad i, j > n. \end{aligned}$$

2. Set $v_N = 1$ and compute for $n = N-1, N-2, \ldots, 0$

$$v_n = \sum_{j=n+1}^{N} a_{nj} v_j.$$

3. Compute for $i = 0, 1, \ldots, N$,

$$p_i = \frac{v_i}{\sum_{j=0}^{N} v_j}.$$

Figure 1: Gaussian elimination

The number of operations required by this algorithm is $O(N^3)$. So it is only practical when N is not too large, say $N \leq 1000$.

We do not apply pivoting, but is this always possible? In other words, in step 1 of the algorithm we divide by a_{nn}; this is feasible if we can prove that $a_{nn} \neq 0$. Further, step 1 involves subtractions, which may cause loss of significant digits. We will show that the algorithm can be slightly adapted such that it only involves multiplication, division and addition of nonnegative numbers. As a consequence, the resulting algorithm is *numerically stable*.

1.2 State space reduction and the GTH algorithm

We first show that Gaussian elimination (without pivoting) has an interesting probabilistic interpretation. Let us consider the coefficients $a_{ij}^{(1)}$ obtained in the first step of this procedure. They satisfy

$$a_{0j}^{(1)} = -\frac{a_{0j}^{(0)}}{a_{00}^{(0)}} = \frac{p_{j0}}{1-p_{00}}\,, \qquad j = 1, \ldots, N,$$

and

$$\begin{aligned} a_{ij}^{(1)} &= a_{ij}^{(0)} + a_{i0}^{(0)} a_{0j}^{(1)} \\ &= p_{ji} - \delta_{ji} + p_{0i}\frac{p_{j0}}{1-p_{00}} \\ &= p_{ji} + p_{0i}\frac{p_{j0}}{1-p_{00}} - \delta_{ji} \qquad i, j = 1, \ldots, N. \end{aligned}$$

If we set

$$p_{ji}^{(1)} = p_{ji} + p_{0i}\frac{p_{j0}}{1-p_{00}}\,, \qquad i, j = 1, \ldots, N,$$

then it follows that

$$a_{ij}^{(1)} = p_{ji}^{(1)} - \delta_{ji} \qquad i, j = 1, \ldots, N.$$

The important observation is that $p_{ij}^{(1)}$ can be interpreted as the transition probability from i to j of the Markov chain *restricted to the set of states* $\{1, 2, \ldots, N\}$: it is possible to directly jump from state j to i (with probability p_{ji}), or first via state 0 and then to i (with probability $p_{j0}p_{0i}/(1-p_{00})$). Hence, after the first step, the equations $1, \ldots, N$ correspond to the equilibrium equations of the Markov chain restricted to $\{1, 2, \ldots, N\}$. So Gaussian elimination may be interpreted as *state space reduction*: we successively eliminate state $1, 2, \ldots$ until only state N remains. By induction we get

$$a_{ij}^{(n)} = p_{ji}^{(n)} - \delta_{ji}, \qquad i, j = n, \ldots, N,$$

where the $p_{ij}^{(n)}$'s are the transition probabilities of the Markov chain restricted to the states $\{n, n+1, \ldots, N\}$. Since the original Markov chain is irreducible, it follows that the restricted Markov chain is also irreducible, and thus

$$-a_{nn}^{(n)} = 1 - p_{nn}^{(n)} > 0.$$

Furthermore,

$$-a_{nn}^{(n)} = 1 - p_{nn}^{(n)} = \sum_{j=n+1}^{N} p_{nj}^{(n)} = \sum_{j=n+1}^{N} a_{jn}^{(n)}.$$

So we can replace $a_{nn}^{(n)}$ in (5) by a column sum, yielding

$$a_{nj}^{(n+1)} = \frac{a_{nj}^{(n)}}{\sum_{k=n+1}^{N} a_{kn}^{(n)}}\,, \qquad j > n,$$

and in (6) we can omit the calculation of the diagonal coefficients $a_{ii}^{(n+1)}$. The resulting version of Gaussian elimination is summarized in Figure 2; it is known as the algorithm of Grassmann, Taksar and Heyman (GTH), see [1].

Algorithm II (GTH algorithm)

0. Set $a_{ij} = p_{ji} - \delta_{ji}$ for all i and j.

1. Compute for $n = 0, 1, \ldots, N-1$

$$\begin{aligned} a_{nj} &= \frac{a_{nj}}{\sum_{k=n+1}^{N} a_{kn}}, \qquad j > n, \\ a_{ij} &= a_{ij} + a_{in} a_{nj}, \qquad i, j > n, i \neq j. \end{aligned}$$

2. Set $v_N = 1$ and compute for $n = N-1, N-2, \ldots, 0$

$$v_n = \sum_{j=n+1}^{N} a_{nj} v_j.$$

3. Compute for $i = 0, 1, \ldots, N$,

$$p_i = \frac{v_i}{\sum_{j=0}^{N} v_j}.$$

Figure 2: GTH algorithm

As we already mentioned, the Gaussian elimination algorithms are only useful when N is not too large. Many real life applications, however, lead to Markov chains with a very large number of states. But the number of outgoing transitions from each state is usually quite small, i.e., P is a sparse matrix. This property is exploited by iterative methods for the solution of Markov chains; they will be treated in the next chapter.

Remark 1.1
Another way to look at the GTH algorithm is as follows. Denote the transition probabilities of the Markov chain restricted to $\{n, n+1, \ldots, N\}$ by $p_{ij}^{(n)}$, $i, j \geq n$. To eliminate state n, we compute the new transition probabilities $p_{ij}^{(n+1)}$ from

$$p_{ij}^{(n+1)} = p_{ij}^{(n)} + p_{in}^{(n)} \cdot \frac{p_{nj}^{(n)}}{\sum_{k=n+1}^{N} p_{nk}^{(n)}}, \qquad i, j > n, i \neq j.$$

and store the transition probabilities from and to state n, i.e., $p_{nj}^{(n)}$ and $p_{in}^{(n)}$. Then, afterwards, we retrieve p_n from the balance equation in state n,

$$p_n \sum_{j=n+1}^{N} p_{nj}^{(n)} = \sum_{i=n+1}^{N} p_i p_{in}^{(n)}.$$

The resulting algorithm is summarized below; it is just a reformulation of the algorithm in Figure 2.

Algorithm III (GTH algorithm)

0. Set p_{ij} to the transition probabilities of the Markov chain on $\{0, 1, \ldots, N\}$.

1. Compute for $n = 0, 1, \ldots, N-1$

$$p_{ij} = p_{ij} + p_{in} \cdot \frac{p_{nj}}{\sum_{k=n+1}^{N} p_{nk}}, \qquad i, j > n, i \neq j.$$

2. Set $v_N = 1$ and compute for $n = N-1, N-2, \ldots, 0$

$$v_n = \frac{1}{\sum_{j=n+1}^{N} p_{nj}} \sum_{j=n+1}^{N} v_j p_{jn}.$$

3. Compute for $i = 0, 1, \ldots, N$,

$$p_i = \frac{v_i}{\sum_{j=0}^{N} v_j}.$$

Figure 3: GTH algorithm

1.3 Exercises

Exercise 1.
In a communication network data moves from switch to switch in the form of packets (strings of bits) of constant length. We may think of a switch as a storage device where packets arrive according to some random process. They are stored in a buffer with a capacity of say K packets and are transmitted one by one. Time is slotted into intervals of fixed length (say one micro second). If at the beginning of a slot, there is a packet in the buffer, it is removed (transmitted) instantaneously; if there is no packet, nothing is done, even if more packets arrive during the slot. If a packet arrives and the buffer is full, it is lost.

(i) Model the switch as a Markov chain.

(ii) Write down the equilibrium equations.

(iii) Show how performance characteristics such as, e.g., mean buffer level and fraction of packets that is lost, can be expressed in terms of the equilibrium probabilities.

Exercise 2.
Now consider a switch with several finite capacity input buffers (for storage of packets from different links). In each slot a packet arrives in buffer i with probability p_i, and exactly one packet can be removed (transmitted) from one of the buffers. There are several service policies for selecting the next packet to be transmitted; e.g., random selection, cyclic selection, or from the buffer with the largest (smallest) number of packets. An important performance characteristic is the number of packets that is lost per time unit. For several service policies,

(i) model the switch as a Markov chain;

(ii) formulate the equilibrium equations;

(iii) and express the fraction of packets that is lost in terms of the equilibrium probabilities.

References

[1] W.K. Grassmann, M.I. Taksar, D.P. Heyman, *Regenerative analysis and steady state distributions for Markov chains*, Opns. Res., 33 (1985), pp. 1107–1116.

[2] W.J. Stewart, *Introduction to the numerical solution of Markov chains*, Princeton University Press, Princeton, 1994.

[3] J. Stoer, R. Bulirsch, *Introduction to numerical analysis*, Springer-Verlag, New York, 1980.