Modeling via Recurrence Relations [1]

Wai-Ki CHING

Department of Mathematics

The University of Hong Kong

1. Mathematical Modeling.
2. Some Problems and Motivations.
3. Some Mathematical Treatment for Second-order Homogeneous Difference Equations.
4. A Rabbit Population Problem.
5. An Eigenvalue Problem.
6. Systems of Difference Equations.
   -An Epidemic Model: The SIR Model
   -An Economic Model: The Cobweb Model
   -A Transportation Model

[1] Available at https://hkumath.hku.hk/~wkc/talks/modeling.pdf

1 Mathematical Modeling

• A mathematical model [2] is a description of a **system** using **mathematical concepts** and **language**. The **process** of developing a mathematical model is termed **mathematical modeling**.

• Mathematical models are used in natural **Sciences** (such as Physics, Chemistry, Biology, Earth Science) and **Engineering** and **Technology** (such as Computer Science and Artificial Intelligence), as well as **Social Sciences** (such as Economics, Psychology, Sociology, Political Science), and **Management**.

• Physicists, engineers, statisticians, operations research analysts, and economists use mathematical models extensively. A model helps to **explain** a system, to study the effects of different components, and to make **predictions about behavior** and rational **decisions**.

[2] Taken from Wikipedia, the free encyclopedia. (http://en.wikipedia.org/wiki/Mathematical_model)

• Mathematical models can take many forms, including but not limited to **Dynamical Systems**, **Statistical Models**, **Differential/** $\boxed{\text{DifferenceEquations}}$, etc.

• In general, mathematical models may include logical models. The quality of a scientific field depends on how well the mathematical models developed on the theoretical side **agree** with the results of repeatable experiments. Lack of agreement between **theoretical mathematical models** and **experimental measurements** often leads to important advances as better theories are developed.
-All models are **wrong**, but some are useful (George E. P. Box).
-Any intelligent fool can make things bigger and more complex. It takes a touch of genius and a lot of courage to move in the opposite direction (Albert Einstein)
- **KISS Principle**, Keep It Simple and Smart.

2 Some Problems and Motivations

2.1 A Counting Problem

- How many different ways of adding up 1's and 2's to 14?

$$
\begin{array}{ll}
1 = 1 & T_1 = 1 \\
2 = 1+1, 2 & T_2 = 2 \\
3 = 1+1+1, 1+2, 2+1 & T_3 = 3 \\
4 = 1+1+1+1, 1+1+2, 1+2+1, 2+1+1, 2+2 & T_4 = 5 \\
\vdots & \vdots
\end{array}
$$

- We let T_j be the number of ways of adding 1's and 2's to j.

- We may have the following relation (**Why?**):

$$\boxed{T_j = T_{j-1} + T_{j-2}.}$$

- This is called an **Initial Value Problem (IVP)**.

- Then you can compute T_{14} as you know T_1 and T_2 (Exercise).

$T(1) = 1 \quad \text{and} \quad T(2) = 2,$
For $j = 3 : 14$,
$T(j) = T(j-1) + T(j-2)$;
end;
Output $T(14)$;

Or using EXCEL:

`https://hkumath.hku.hk/~wkc/talks/pattern1.xlsx`

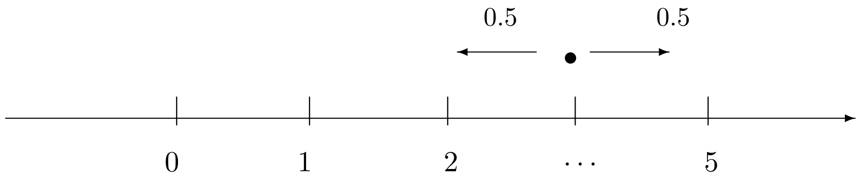

Figure 1: The random walk.

2.2 A Probability Problem

• A walker performs a random walk on the grid points from 0 to 5. Each time the walker tosses a fair coin, the walker moves one step forward if it is a head otherwise the walker moves one step backward. The walk stops when the random walker is at position 0 or 5.

• What is the probability that the walker ends up at 0 given that he begins at j $(j = 1, 2, 3, 4)$?

- We let the probability be T_j.

Then we have $T_0 = 1$ and $T_5 = 0$.

Moreover, for $j = 1, 2, 3, 4$, we have

$$\boxed{T_j = 0.5 \cdot T_{j+1} + 0.5 \cdot T_{j-1}.}$$

- This is called a **Boundary Value Problem (BVP)** (Exercise).

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & -1 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & -1 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & -1 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & -1 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.8000 & -1.6000 & -1.2000 & -0.8000 & -0.4000 & 0.2000 \\ 0.6000 & -1.2000 & -2.4000 & -1.6000 & -0.8000 & 0.4000 \\ 0.4000 & -0.8000 & -1.6000 & -2.4000 & -1.2000 & 0.6000 \\ 0.2000 & -0.4000 & -0.8000 & -1.2000 & -1.6000 & 0.8000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 0.8 \\ 0.6 \\ 0.4 \\ 0.2 \\ 0.0 \end{bmatrix}$$

3 Some Mathematical Treatment for Second-order Homogeneous Difference Equations

• The previous problems are **second-order homogeneous difference equation** of the following form:

$$\boxed{a_2T_{j+2} + a_1T_{j+1} + a_0T_j = 0, \quad j = 0, 1, 2, \ldots,} \tag{3.1}$$

• We seek for solution of the form:

$$\boxed{T_j = z^j.}$$

If we substitute it into Eq. (3.1), we get the **Euler equation**:

$$\boxed{f(z) = a_2z^2 + a_1z + a_0 = 0.}$$

• The roots are given by

$$z_1 = \frac{-a_1 - \sqrt{a_1^2 - 4a_0a_2}}{2a_2}$$

and

$$z_2 = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_2}.$$

Case 1: Suppose that $z_1 \neq z_2$ and both are real numbers.

Then both

$$\boxed{T_j = c_1 \cdot z_1^j \quad \text{and} \quad T_j = c_2 \cdot z_2^j}$$

are solutions to Eq. (3.1) for any real numbers c_1 and c_2.

• We can also check that their sum

$$\boxed{T_j = c_1 \cdot z_1^j + c_2 \cdot z_2^j, \quad j = 0, 1, \ldots,} \tag{3.2}$$

is also a solution. Therefore it is the **general solution** to Eq. (3.1).

Case 2: If $z_1 = z_2$ (repeated root), then apart from the solution

$$\boxed{T_j = c_1 \cdot z_1^j}$$

another solution will be

$$\boxed{T_j = c_2 \cdot j \cdot z_1^j.}$$

Because when we substitute $T_j = j \cdot z_1^j$ into Eq. (3.1), we get

$$z_1^j \cdot \left(j \cdot (a_2 \cdot z_1^2 + a_1 \cdot z_1 + a_0) + (2 \cdot a_2 \cdot z_1^2 + a_1 \cdot z_1)\right)$$

and

$$f(z_1) = a_2 \cdot z_1^2 + a_1 \cdot z_1 + a_0 = 0$$

and

$$z_1 \cdot (2 \cdot a_2 \cdot z_1 + a_1) = z_1 f'(z_1) = 0.$$

- Therefore the general solution is of the form:

$$\boxed{T_j = z_1^j \cdot (c_1 + c_2 j), \quad j = 0, 1, \ldots,} \tag{3.3}$$

Case 3: Complex Roots. In this case, we consider the difference equation of the following form (where $a, b \in \mathbb{R}$):

$$\boxed{T_{j+2} - 2aT_{j+1} + (a^2 + b^2)T_j = 0.}$$

- Then $a + bi$ and $a - bi$ form a pair of complex (conjugate) roots of the equation:

$$x^2 - 2ax + (a^2 + b^2) = 0.$$

Then we have the solutions in their **polar form**:

$$a + bi = \rho(\cos\theta + i\sin\theta)$$

and

$$a - bi = \rho(\cos\theta - i\sin\theta)$$

where

$$\rho = \sqrt{a^2 + b^2} \quad \text{and} \quad \tan\theta = \left|\frac{b}{a}\right|. \tag{3.4}$$

• By **de Moivre's Theorem** we have for $j \in \mathbb{N}$

$$(a+bi)^j = \rho^j(\cos\theta + i\sin\theta)^j = \rho^j(\cos j\theta + i\sin j\theta)$$

and

$$(a-bi)^j = \rho^j(\cos\theta - i\sin\theta)^j = \rho^j(\cos j\theta - i\sin j\theta)$$

which can be proved by using mathematical induction on n (exercise).

• Thus

$$\begin{aligned} A_1(a+bi)^j + A_2(a-bi)^j &= A_1\rho^j(\cos j\theta + i\sin j\theta) + A_2\rho^j(\cos j\theta - i\sin j\theta) \\ &= (A_1+A_2)\rho^j\cos j\theta + i(A_1-A_2)\rho^j\sin j\theta. \end{aligned}$$

• Finally, in all the cases, the parameters (unknowns) can be solved if we are given extra information such as T_0 and T_1, the initial conditions.

Exercise: $T_{j+2} - T_{j+1} + T_j = 0$ with $T_0 = 0$ and $T_1 = 1$.

4 A Rabbit Population Problem

• One of the early examples of a recursively defined sequence arises in the writings of **Fibonacci**, who was the greatest European Mathematician in the Middle Ages.

• In 1202, Fibonacci posed the following problem about the number of rabbits in a closed environment. He assumed that a single pair of rabbits is born at the beginning of a year. We then assume the following conditions:

(i) Rabbit pairs are **not fertile** during their first month of life but thereafter give birth to one new male/female pair at the end of every month.

(ii) No death occurs during the year

• The question is: How many rabbits will there be at the end of the year?

• Let T_j be the number of rabbit pairs alive at the end of month j. $T_0 = 1$, the initial number of rabbit pairs. We also have $T_1 = 1$

• The number of rabbit pairs at the end of month j is equal to the sum of **the number of rabbit pairs at the end of month** $j-1$ and **the number of rabbit pairs born at the end of month** j.

• **The number of rabbit pairs born at the end of month j is equal to the number of rabbit pairs at the end of month** $j-2$.

• The recurrence relation $(j = 2, 3, \ldots)$ is then given by

$$\boxed{T_j = T_{j-1} + T_{j-2} \quad \text{with} \quad T_0 = T_1 = 1.}$$

Therefore the answer for the Fibonacci's question is $T_{12} = 233$. How?

• In fact, using our techniques in Section 3, we can obtain (Exercise):

$$\boxed{T_j = \frac{\sqrt{5}+1}{2\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^j + \frac{\sqrt{5}-1}{2\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^j.}$$

5 An Eigenvalue Problem

- To obtain the eigenvalues and eigenvectors of a special $n \times n$ matrix:

$$A_n = \begin{bmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ 0 & & & & -1 & 2 \end{bmatrix}$$

- We consider the relation $A_n\mathbf{v} = \lambda\mathbf{v}$, then we have

$$\begin{cases} (2-\lambda)v_1 - v_2 & = & 0 \\ -v_1 + (2-\lambda)v_2 - v_3 & = & 0 \\ \vdots & & \vdots \\ -v_{j-1} + (2-\lambda)v_j - v_{j+1} & = & 0 \\ \vdots & & \vdots \\ -v_{n-1} + (2-\lambda)v_n & = & 0 \end{cases}$$

where $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T \neq \mathbf{0}$ is **an eigenvector**.

- We define $v_0 = v_{n+1} = 0$ then we have the difference equations:

$$-v_{j-1} + (2-\lambda)v_j - v_{j+1} = 0 \quad \text{for} \quad j = 1, 2, \ldots, n.$$

or

$$\boxed{v_{j+1} - (2-\lambda)v_j + v_{j-1} = 0 \quad \text{for} \quad j = 1, 2, \ldots, n.}$$

- Using the result we learned, the solution is of the form:

$$\boxed{v_j = Bm_1^j + Cm_2^j}$$

where B and C are two constants and m_1 and m_2 are roots of the quadratic equation

$$\boxed{m^2 - (2-\lambda)m + 1 = 0} \tag{5.1}$$

- We remark that $m_1 \neq m_2$. Because if this is the case, we have the solution form:

$$v_j = (B + Cj)m_1^j$$

but $v_0 = v_{n+1} = 0$ implies that

$$v_0 = 0 = B \quad \text{and} \quad v_{n+1} = 0 = C(n+1)m_1^{n+1}$$

and therefore $B = C = 0$ as $m_1 \neq 0$.

Now we have

$$v_0 = 0 = B + C$$

and

$$v_{n+1} = 0 = Bm_1^{n+1} + Cm_2^{n+1}.$$

Thus we have

$$B = -C \quad \text{and} \quad \left(\frac{m_1}{m_2}\right)^{n+1} = 1 = e^{2\pi i}$$

where $i = \sqrt{-1}$. Thus

$$\frac{m_1}{m_2} = e^{\frac{2s\pi i}{n+1}} \quad \text{for} \quad s = 1, 2, \ldots, n.$$

- By Eq. (5.1), we have

$$m_1 m_2 = 1.$$

Here we have

$$m_1 = e^{\frac{s\pi i}{n+1}} \quad \text{and} \quad m_2 = e^{-\frac{s\pi i}{n+1}}$$

- By Eq. (5.1) again, we have

$$m_1 + m_2 = 2 - \lambda$$

and thus we have for $s = 1, 2, \ldots, n$,

$$\begin{aligned}
\lambda_s &= 2 - (e^{\frac{s\pi i}{n+1}} + e^{\frac{-s\pi i}{n+1}}) \\
&= 2 - 2\cos(\tfrac{s\pi}{n+1}) \qquad (\cos(2\theta) = 1 - 2\sin^2(\theta)) \\
&= 2 - 2(1 - 2\sin^2(\tfrac{s\pi}{2(n+1)})) \\
&= 4\sin^2(\tfrac{s\pi}{2(n+1)}).
\end{aligned}$$

Finally we note that for λ_s, we have $j = 1, 2, \ldots, n$

$$\boxed{v_j = Bm_1^j + Cm_2^j = B(e^{\frac{s\pi i}{n+1}} - e^{\frac{-s\pi i}{n+1}}) = 2iB\sin\left(\frac{js\pi}{n+1}\right).}$$

- As $2iB$ is just a constant, we have

$$\boxed{\mathbf{v}_s = \left[\sin\left(\frac{s\pi}{n+1}\right), \sin\left(\frac{2s\pi}{n+1}\right), \cdots, \sin\left(\frac{sn\pi}{n+1}\right)\right]^T.}$$

6 Systems of Difference Equations

6.1 An Epidemic Model: The SIR Model

Modeling the spread of an epidemic was issued by Kermack and McKendrick (1927). In a closed population of size N, at time t, we let

x_t be the population of the susceptible;

y_t be the population of the infective;

z_t be the population of the removal.

• The total population is assumed to be homogeneously mixed and is constant at any time t, i.e., $x_t + y_t + z_t = N$.

• Let β **be the infection rate** and γ **be the removal rate**. Decrease rate of $x_{t+1} \propto x_t$ and y_t. Then the system dynamics can be described as follows:

$$\begin{cases} x_{t+1} &= x_t - \beta \cdot x_t \cdot y_t \\ y_{t+1} &= y_t + \beta \cdot x_t \cdot y_t - \gamma \cdot y_t \\ z_{t+1} &= z_t + \gamma \cdot y_t \end{cases}$$

- Suppose $\beta = 0.02$ and $\gamma = 0.2$.

- The population size $N = 100$ and initial at $t = 0$ there is one infective. Using the difference equations, we know within one week, all people will be infected.

t	x_t	y_t	z_t
0	99.0	1.0	0.0
1	97.0	2.8	0.2
2	91.6	7.6	0.8
3	77.7	20.1	2.2
4	46.5	47.2	6.3
5	2.6	81.7	15.8
6	0.0	69.6	30.4

6.2 An Economic Model: The Cobweb Model.

• The Cobweb model is an economic theory for studying price fluctuation and market equilibrium. Let p_t be the price, q_t be quantity at time t, and the model reads:

$$\begin{cases} q_t &= \alpha + \beta p_{t-1} \\ p_t &= \gamma - \delta q_t \end{cases}$$

where where $\alpha, \beta > 0, \gamma, \delta > 0$ and $q_0 > 0$ are given and therefore $p_0 = \gamma - \delta q_0$.

• Solving the above two equations, we obtain two first-order linear difference equations:

$$q_t = \alpha + \beta\gamma - \beta\delta q_{t-1}$$

and

$$p_t = \gamma - \alpha\delta - \beta\delta p_{t-1}.$$

It is straightforward to solve for p_t and q_t, but the long-run behavior of the solutions are of special interest.

• Since $\beta > 0, \delta > 0$, their product $\beta\delta > 0$, the solutions are thus always oscillatory with respect to time t.

• As $t \to \infty$, the equilibrium point is

$$\boxed{(p^*, q^*) = \Big(\frac{\gamma - \alpha\delta}{1+\beta\delta}, \frac{\alpha+\beta\gamma}{1+\beta\delta}\Big)}$$

which is obtained by solving

$$\begin{cases} q = \alpha + \beta\gamma - \beta\delta q \\ p = \gamma - \alpha\delta - \beta\delta p. \end{cases}$$

(1) If $0 < \beta\delta < 1$, then the solutions $\{p_t\}$ and $\{q_t\}$ will converge to (p^*, q^*).

(2) If $\beta\delta = 1$, the solutions will oscillate with finite magnitude.

(3) If $\beta\delta > 1$, the solutions will oscillate with infinite magnitude.

6.3 A Transportation Model

• Consider a random walker (traveler) in a network of cities as shown in Figure 2. The walker travels every day to the cities in the network.

(i) Suppose that at City i ($i = 1, 2, 3, 4, 5$), the probabilities of traveling to other adjacent cities are all equal.

(ii) While the probability of staying at the same cite in the next move is assumed to be zero.

(iii) At time 0 (day 0), the traveler is in **City** 2.

• **Question:** What is the probability that at **time** 7 **(day** 7**)** (one week later), the traveler is found in **City** 1**?**

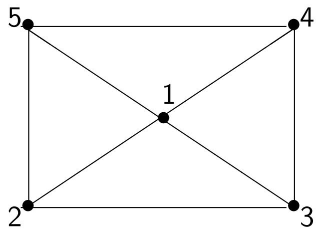

Figure 2: The random walker and the network.

Observations:

We note that the probability of which city to visit tomorrow (time $t+1$) will be fixed if the current city that the traveler visited (time t) is known.

Suppose today the traveler is in City 1 (time t), then with probability $1/4$ he will visit Cities 2,3,4 or 5 tomorrow (time $t+1$).

If today the traveler is in City 2 (time t), then with probability $1/3$ he will visit Cities 1,3,or 5 tomorrow (time $t+1$). But the probability of visiting City 4 is 0.

- **Problem Formulation:**

We first define the **transition probability** Q_{ij}. Here Q_{ij} is the probability that the traveler will make a move to City i given the traveler is now in City j.

$Q_{i1} = 1/4$ for $i = 2, 3, 4, 5$;
$Q_{i2} = 1/3$ for $i = 1, 3, 5$; $\quad Q_{i3} = 1/3$ for $i = 1, 2, 4$;
$Q_{i4} = 1/3$ for $i = 1, 3, 5$; $\quad Q_{i5} = 1/3$ for $i = 1, 2, 4$;

The remaining Q_{ij} are all zero. If we know the probabilities, then we can put them into a matrix:

$$Q = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}.$$

Let $P_i(t)$ be the probability that the traveler is in City i at time t. Here $i = 1, 2, 3, 4, 5$ and $t = 0, 1, 2, \ldots,$.

Now we have

$$\begin{cases} P_1(t+1) = & & Q_{12}P_2(t) & +Q_{13}P_3(t) & +Q_{14}P_4(t) & +Q_{15}P_5(t) \\ P_2(t+1) = & Q_{21}P_1(t) & & +Q_{23}P_3(t) & & +Q_{25}P_5(t) \\ P_3(t+1) = & Q_{31}P_1(t) & +Q_{32}P_2(t) & & +Q_{34}P_4(t) & \\ P_4(t+1) = & Q_{41}P_1(t) & & Q_{43}P_3(t) & & Q_{45}P_5(t) \\ P_5(t+1) = & Q_{51}P_1(t) & +Q_{52}P_2(t) & & +Q_{54}P_4(t) & \end{cases}$$

and

$$P_1(0) = P_3(0) = P_4(0) = P_5(0) = 0 \quad \text{and} \quad P_2(0) = 1.$$

Or

$$\begin{cases} P_1(t+1) = & & 1/3P_2(t) & +1/3P_3(t) & +1/3P_4(t) & +1/3P_5(t) \\ P_2(t+1) = & 1/4P_1(t) & & +1/3P_3(t) & & +1/3P_5(t) \\ P_3(t+1) = & 1/4P_1(t) & +1/3P_2(t) & & +1/3P_4(t) & \\ P_4(t+1) = & 1/4P_1(t) & & +1/3P_3(t) & & +1/3P_5(t) \\ P_5(t+1) = & 1/4P_1(t) & +1/3P_2(t) & & +1/3P_4(t) & \end{cases}$$

and

$$P_1(0) = P_3(0) = P_4(0) = P_5(0) = 0 \quad \text{and} \quad P_2(0) = 1.$$

and we are asked to find $P_1(7)$.

- In matrix language, the relation can be written as follows:

$$\begin{bmatrix} P_1(t+1) \\ P_2(t+1) \\ P_3(t+1) \\ P_4(t+1) \\ P_5(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \\ P_5(t) \end{bmatrix} \tag{6.1}$$

- If we let

$$\mathbf{P}(t) = [P_1(t)\ P_2(t)\ P_3(t)\ P_4(t)\ P_5(t)]^T$$

then we may write

$$\boxed{\mathbf{P}(t+1) = Q\mathbf{P}(t)}$$

and

$$\mathbf{P}(0) = [P_1(0)\ P_2(0)\ P_3(0)\ P_4(0)\ P_5(0)]^T = [0\ 1\ 0\ 0\ 0]^T$$

and we are asked to find $P_1(7)$ of $\mathbf{P}(7)$.

• Using EXCEL/SciLab/MatLab, we compute the followings:

https://hkumath.hku.hk/~wkc/talks/traveler1.xlsx

$\mathbf{P(1)}$	$\mathbf{P(2)}$	$\mathbf{P(3)}$	$\mathbf{P(4)}$	$\mathbf{P(5)}$	$\mathbf{P(6)}$	$\mathbf{P(7)}$
0.3333	0.2222	0.2593	0.2469	0.2510	0.2497	**0.2501**
0.0000	0.3056	0.1111	0.2376	0.1543	0.2095	0.1728
0.3333	0.0833	0.2593	0.1389	0.2202	0.1656	0.2021
0.0000	0.3056	0.1111	0.2377	0.1543	0.2095	0.1728
0.3333	0.0833	0.2593	0.1389	0.2202	0.1656	0.2021

The answer is then equal to 0.2501 roughly.

Remark 1: One may consider the long-run probabilities, i.e.,

$$\lim_{t\to\infty} P_i(t) = P_i, \quad i = 1, 2, 3, 4, 5.$$

Suppose they exist, then they must satisfy

$$\begin{cases} P_1 = & & 1/3P_2 & +1/3P_3 & +1/3P_4 & +1/3P_5 \\ P_2 = & 1/4P_1 & & +1/3P_3 & & +1/3P_5 \\ P_3 = & 1/4P_1 & +1/3P_2 & & +1/3P_4 & \\ P_4 = & 1/4P_1 & & +1/3P_3 & & +1/3P_5 \\ P_5 = & 1/4P_1 & +1/3P_2 & & +1/3P_4 & \end{cases} \tag{6.2}$$

and

$$P_1 + P_2 + P_3 + P_4 + P_5 = 1.$$

• What is the meaning of P_i? It means when the traveling process has been continued for a long time (**in equilibrium**), the probability of finding the traveler in City i will be P_i

• Here by symmetry of the network, $P_2 = P_3 = P_4 = P_5 = P$, we have

$$P_1 + 4P = 1 \quad \text{and} \quad P_1 = 4/3P \ \ (\text{the first equation}).$$

Hence $P_1 = 1/4$ and $P_2 = P_3 = P_4 = P_5 = P = 3/16$.

$$\lim_{t\to\infty} \mathbf{P}(t) = \left[\frac{4}{16}\ \frac{3}{16}\ \frac{3}{16}\ \frac{3}{16}\ \frac{3}{16}\right]^T.$$

• The traveler is "spending" 25% of his time in city 1. While the other four cities will share equally the remaining 75% of his visiting time.

• If we regard a city as a webpage, the traveler as a user on the internet and the edges as the links. Then webpage 1 is the most important one (rank no. 1) because the user spends most of his time (25%) on this webpage. The other four webpages have the same ranking. This is the key idea of Google's PageRank algorithm.

Remark 2: In the example, we ignore the **transportation cost** which may affect the traveler's decision. Suppose the highways connecting the cities have different fees (costs) given in the following table.

City	1	2	3	4	5
1	∞	20	25	40	50
2	20	∞	20	∞	40
3	25	20	∞	25	∞
4	40	∞	25	∞	10
5	50	40	∞	10	∞

Question: How to model the probabilities in the matrix of Eq. (6.1)?

Let Q_{ij} be the probability that the traveler will move to City i tomorrow given that today he is in City j. Let C_{ij} be the cost of taking the highway from City i to City j (in the above table). What is an appropriate relation between Q_{ij} and C_{ij}?

• One possible and reasonable suggestion is to assume that Q_{ij} is a decreasing function in C_{ij}. For example,

(a) $Q_{ij} \propto \frac{1}{C_{ij}}$ [3]

(b) $Q_{ij} \propto e^{-C_{ij}}$

Suppose we adopt (a), then let us compute Q_{i1} for $i = 1, 2, 3, 4$. We have (for some $K > 0$),

$$Q_{21} = \frac{K}{20}, \; Q_{31} = \frac{K}{25}, \; Q_{41} = \frac{K}{40}, \; Q_{51} = \frac{K}{50}.$$

To determine K, we note that (why?)

$$Q_{21} + Q_{31} + Q_{41} + Q_{51} = 1.$$

Hence we have

$$K(0.2 + 0.25 + +0.025 + 0.02) = 1 \quad \text{or} \quad K = 7.4074.$$

[3] You may consider $Q_{ij} \propto \frac{1}{C_{ij}+1}$ if C_{ij} is too close to 0

- We can compute the remaining probabilities similarly.

- We have a new system as follows:

$$\begin{bmatrix} P_1(t+1) \\ P_2(t+1) \\ P_3(t+1) \\ P_4(t+1) \\ P_5(t+1) \end{bmatrix} = \begin{bmatrix} 0.0000 & 0.4000 & 0.3077 & 0.1515 & 0.1379 \\ 0.3704 & 0.0000 & 0.3846 & 0.0000 & 0.1724 \\ 0.2963 & 0.4000 & 0.0000 & 0.2424 & 0.0000 \\ 0.1852 & 0.0000 & 0.3077 & 0.0000 & 0.6897 \\ 0.1481 & 0.2000 & 0.0000 & 0.6061 & 0.0000 \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \\ P_5(t) \end{bmatrix}$$

In this case, we have

$$\mathbf{P}(7) = [0.1971\ 0.1557\ 0.2104\ 0.1890\ 0.2477]^T.$$

and $P_1(7) = 0.1971$.

- By solving the linear system of equations similar to Eq. (6.2), we can show that

$$\lim_{t\to\infty} \mathbf{P}(t) = [0.1929\ 0.1786\ 0.1857\ 0.2357\ 0.2071]^T.$$

7 References

1. S. Elaydi (2005) *An Introduction to Difference Equations*, Third Edition, Springer, New York.
2. S. Goldberg (2010) *Introduction to Difference Equations*, Dover Publications Inc., New York.
3. H. Tijms (2012) *Understanding Probability*, Third Edition, Cambridge University Press, Cambridge.
4. W. Kermack and A. McKendrick (1927) A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115, 700–721. `https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1927.0118`
5. J. Ortuzar and L. Willumsen (2011) *Modelling Transport*, Fourth Edition, Wiley, New York.