MTL-106
Probability and Stochastic Processes

Assignment - 2D

Deadline : 10th April 2024

To illustrate the application of transition matrices in predicting the spread of a disease in a population, let's consider a simplified scenario. Transition matrices are powerful tools in modeling various processes, including epidemiology, where they can help predict the future state of a system based on its current state.

1 Scenario Description

Imagine we have a small population divided into three categories based on their health status in relation to a particular disease:

- **S (Susceptible)**: Individuals who are healthy but can catch the disease.
- **I (Infected)**: Individuals who have caught the disease and can spread it.
- **R (Recovered)**: Individuals who have recovered from the disease and are now immune.

We assume that over a fixed time interval (e.g., a day or a week), individuals can transition between these states according to certain probabilities. These transitions are summarized in a transition matrix, where each element a_{ij} represents the probability of moving from state i to state j in one time step.

2 Transition Matrix

Let's denote our states as follows: $S = 1, I = 2, R = 3$. Our transition matrix A might look something like this:

$$A = \begin{pmatrix} P_{SS} & P_{SI} & P_{SR} \\ P_{IS} & P_{II} & P_{IR} \\ P_{RS} & P_{RI} & P_{RR} \end{pmatrix}$$

- P_{SS}, P_{II}, P_{RR} are the probabilities of staying in the same state for Susceptible, Infected, and Recovered, respectively. - P_{SI} is the probability of a Susceptible individual becoming Infected. - P_{SR} is essentially 0, as Susceptible individuals cannot directly become Recovered without being Infected first. - P_{IS} is also 0, assuming infected individuals cannot become susceptible again without recovering first. - P_{IR} is the probability of an Infected individual Recovering. - P_{RS} and P_{RI} are 0, assuming recovered individuals are immune and cannot become Susceptible or Infected again. - P_{RR} is 1, assuming recovered individuals stay recovered (assuming lifelong immunity for simplification).

3 Example Matrix

Suppose we have a transition matrix as follows:

$$A = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0 & 0.8 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$$

This matrix indicates:

- 10% of Susceptible individuals become Infected each time step.
- 20% of Infected individuals Recover each time step.
- Recovered individuals remain Recovered.

4 Predicting Disease Spread

Given an initial state vector $v_0 = (S_0, I_0, R_0)$, representing the initial number (or proportion) of individuals in each state, we can predict the state vector v_n after n time steps by applying the transition matrix n times:

$$v_n = A^n v_0$$

This calculation will give us the predicted distribution of the population across the three health statuses after n time steps.

Let's expand the scenario to introduce more complexity and realism into our model of disease spread using transition matrices. We'll include additional states to account for exposed individuals and differentiate between symptomatic and asymptomatic infections. Furthermore, we'll consider the impact of vaccination and natural death or birth rates, making our model more applicable to real-world scenarios.

5 Expanded Scenario Description

Our population is now divided into the following categories:

- **S (Susceptible)**: Individuals who are healthy but can catch the disease.
- **E (Exposed)**: Individuals who have been exposed to the disease but are not yet infectious.
- **I (Infected)**: Individuals who are infected and can spread the disease. This category is further divided into:
- **Ia (Asymptomatic Infected)**: Infected individuals who do not show symptoms but can spread the disease.
- **Is (Symptomatic Infected)**: Infected individuals who show symptoms and can spread the disease.
- **R (Recovered)**: Individuals who have recovered from the disease and are now immune.
- **V (Vaccinated)**: Individuals who have been vaccinated and are assumed to be immune.
- **D (Deceased)**: Individuals who have died from the disease.

6 Extended Transition Matrix

With these categories, our transition matrix A becomes significantly larger, reflecting the complexity of disease dynamics:

$$A = \begin{pmatrix} P_{SS} & P_{SE} & 0 & 0 & 0 & 0 & 0 \\ 0 & P_{EE} & P_{EIa} & P_{EIs} & 0 & 0 & 0 \\ 0 & 0 & P_{IaIa} & 0 & P_{IaR} & 0 & P_{IaD} \\ 0 & 0 & 0 & P_{IsIs} & P_{IsR} & 0 & P_{IsD} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ P_{VS} & 0 & 0 & 0 & 0 & P_{VV} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

This matrix represents a more detailed set of transitions: - Susceptible individuals can become Exposed (P_{SE}). - Exposed individuals progress to either Asymptomatic or Symptomatic Infected states (P_{EIa}, P_{EIs}). - Infected individuals can either recover (P_{IaR}, P_{IsR}) or, unfortunately, die (P_{IaD}, P_{IsD}). - Vaccinated individuals might return to the Susceptible category if vaccine immunity wanes (P_{VS}) or remain in the Vaccinated state (P_{VV}), with a small chance of moving to the Recovered state (P_{VR}) if exposed and the vaccine provides sterilizing immunity. - Natural deaths and births could be represented in this model but to keep the focus on disease dynamics, these aspects are simplified here.

Now, let's add some more dynamics to model things like vaccination and social distancing

7 Add More Dynamics

- **Vaccination ($P_{SV}(t)$)**: The probability that a Susceptible individual becomes Vaccinated at time t. This rate could increase following the start of a vaccination campaign at t_0.

$$P_{SV}(t) = \begin{cases} 0 & \text{if } t < t_0 \\ \alpha(t - t_0) & \text{if } t \geq t_0 \end{cases}$$

  Where α is a parameter representing the vaccination rate, which could itself vary over time based on supply and uptake.
  As the value of P_{SV} is initially zero and increases with time, to ensure that the total transition probability from state S to any other state remains 1, we will decrease the P_{SS} and P_{SE} in the following manner:

$$P_{SS}(t) = \frac{(1 - P_{SV}(t)) * P_{SS}(t-1)}{P_{SE}(t-1) + P_{SS}(t-1)}$$

$$P_{SE}(t) = \frac{(1 - P_{SV}(t)) * P_{SE}(t-1)}{P_{SE}(t-1) + P_{SS}(t-1)}$$

- **Social Distancing ($P_{EI}(t)$)**: The probability that an Exposed individual becomes Infected (either asymptomatic or symptomatic). This probability decreases during periods of enhanced social distancing measures.

  Suppose social distancing measures are enhanced between t_1 and t_2, then:

$$P_{EI}(t) = \begin{cases} \beta & \text{if } t < t_1 \text{ or } t > t_2 \\ \gamma \cdot \beta & \text{if } t_1 \leq t \leq t_2 \end{cases}$$

  Where β is the base transition rate, and $0 < \gamma < 1$ is a factor that reduces this rate due to social distancing.

- **Seasonal Variation ($P_{SE}(t)$)**: The probability of a Susceptible individual becoming Exposed may vary with seasons. This can be modeled as a periodic function of time, reflecting, for example, higher transmission rates in colder months.

$$P_{SE}(t) = P_{SE}(t-1) + \epsilon \cdot \sin\left(\frac{2\pi}{L} \cdot t\right)$$

$$P_{SS}(t) = P_{SS}(t-1) - \epsilon \cdot \sin\left(\frac{2\pi}{L} \cdot t\right)$$

  Where ϵ represents the amplitude of seasonal variation, L is the period (e.g., 365 days for yearly seasonality).

8 New Model Dynamics

The dynamics of the system can then be described using a set of difference equations or differential equations, depending on whether we approach the model discretely or continuously. For a discrete model, the state at time $t + 1$ is a function of the state at time t, modulated by the transition probabilities:

$$\begin{aligned}
S(t+1) &= S(t) - P_{SE} \cdot S(t) - P_{SV} \cdot S(t) + P_{VS} \cdot V(t) \\
E(t+1) &= E(t) + P_{SE} \cdot S(t) - P_{EI} \cdot E(t) \\
I_a(t+1) &= I_a(t) + \rho \cdot P_{EI} \cdot E(t) - P_{I_aR} \cdot I_a(t) - P_{I_aD} \cdot I_a(t) \\
I_s(t+1) &= I_s(t) + (1-\rho) \cdot P_{EI} \cdot E(t) - P_{I_sR} \cdot I_s(t) - P_{I_sD} \cdot I_s(t) \\
R(t+1) &= R(t) + P_{I_aR} \cdot I_a(t) + P_{I_sR} \cdot I_s(t) \\
V(t+1) &= V(t) + P_{SV} \cdot S(t) \\
D(t+1) &= D(t) + P_{I_sD} \cdot I_s(t) + P_{I_aD} \cdot I_a(t)
\end{aligned}$$

Where $P_{EI} = P_{EIa} + P_{EIs}$ & ρ represents the fraction of Exposed individuals who become Asymptomatic Infected, and $P_{I_aR}(t)$, $P_{I_sR}(t)$, $P_{I_sD}((t)$ are the probabilities of Asymptomatic Infected individuals recovering, Symptomatic Infected individuals recovering, and Symptomatic Infected individuals dying, respectively, each of which may also vary over time due to changes in medical treatment efficiency or healthcare capacity.

This set of equations describes how the number (or proportion) of individuals in each compartment changes over time, influenced by the dynamic transition probabilities that reflect vaccination efforts, public health interventions, and seasonal variations.

9 Coding Part

Your task is to find the distribution of the disease in different categories after a fixed time

9.1 Input

- For each test case, the first 7 lines of the input represent the probability transition matrix as described in the **Section 6** of the assignment.
- The next line contains the values of the parameters t_0, t_1 and t_2
- The next line contains the value for the parameters α, γ, ρ & ϵ

  NOTE: β is simply the initial value of $P_{EI} = P_{EIa} + P_{EIs}$
- The next line contains the time T at which we want the distribution in the population (i.e., at the end of T days)

  You can assume that initially (i.e., at start of day 1), no one is infected and the entire population is in susceptible category, i.e., initially the vector is [1 0 0 0 0 0 0]

NOTE: L is assumed to be 365, and all the time scales are in days, Also, it will be ensured none of the probability values go negative or the total probability goes above 1 over time

9.2 Output

Output 7 terms representing the distribution of disease into different categories after time T in a single line with space separated values.

9.3 Constraints & Time Limits

$T \leq 10^6$
Time-Limit = 3sec

Sample Inputs and Outputs are provided separately (see MSTeams A2 channel), but your code will be tested on a much larger and exhaustive dataset so make sure to handle all the edge cases.

You are expected to write your own code from scratch using only the standard library of C++ or python. Any instances of plagiarism (either among students or copy from the internet) will be awarded 0.