4452 Mathematical Modeling Lecture 16: Markov Processes

Introduction

A stochastic model is one in which random effects are incorporated into the model. The battle simulations of the last lecture were stochastic models.

A Markov chain is a particular type of discrete time stochastic model. A Markov process is a particular type of continuous time stochastic model.

All of what follows should sound very familiar. The lecture on battle simulations essentially took a discrete time system (the battle) and modeled it as deterministic using a discrete dynamical system (recurrence relations), and then stochastically using probabilities and random numbers. This stochastic model was an implementation of a Markov chain, although I did not call it that at the time.

In the continuous dynamical system social mobility example which we will look at in this lecture, we are comparing a deterministic model using differential equations with a stochastic model using a Markov process. We shall see, as we did before, that the stochastic model does not contain time explicitly, and to add time we need to model the time between events as a Poisson process. We choose a Poisson process since it is memoryless–each event has no memory (meaning effect) on the next.

Markov Chains

A Markov chain models a sequence of events, and the probability of an event occurring depends on the previous history of the system.

A Markov chain is a system which starts in one state, and then moves to one of several other states with various probabilities (if the current state is s_i, then there are probabilities that the next state will be s_j).

Definition[1] Markov Chain:

1. States: $s \in \{1, 2, \ldots, m\}$ where m is finite.
2. Starting State: the starting state s_0 may be fixed or drawn from some a priori distribution, $P_0(s_0)$.
3. Dynamics: we define how the system transitions from the current state s_t to the next state s_{t+1}. The transitions satisfy the first order Markov property

$$P(s_{t+1}|s_t, s_{t-1}, \ldots, s_0) = P_1(s_{t+1}|s_t)$$

The resulting stochastic system generates a sequence of states

$$s_0 \to s_1 \to s_2 \to s_3 \to \ldots$$

where s_0 is drawn from $P_0(s_0)$, and s_{t+1} from $P_1(s_{t+1}|s_t)$.

If the probabilities do not depend on time, then the Markov chain is called homogeneous (or, has stationary transition probabilities). In the battle simulation, the probabilities *did* depend on time, so the Markov chain there had non stationary transition probabilities.

The stochastic nature of the Markov chain occurs when the system decides which state to move to next, since this is a random process.

A Markov process is completely defined once its transition probability matrix and initial state are specified.

Markov chains are typically represented in terms of probability matrices, or state diagrams. Hence, there is a link to graph theory. Any search of the web for Markov chains will turn up numerous references to Markov Chain Monte Carlo, Bayesian statistics, statistical physics, graphical methods, random walks, and a host of other applications from genetics to artificial intelligence.

Steady State Distribution

For Markov chains, one talks about the steady state distribution of the probabilities of the Markov chain, and not the steady state as we have thought of it before as in the system stays locked in a single state.

We should always remember that not all Markov processes will have a steady state distribution. If for any pair of states the system can move between the two states in a finite number of transitions, then the Markov process is called *ergodic*. Ergodic Markov processes always tend to a steady state.

The idea of a Markov process is to begin with a state transition probability matrix P,

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{1m} & \cdots & p_{mm} \end{pmatrix}$$

for the m states $\{1, 2, \ldots, m\}$. We have the probabilities

$$\pi_n(i) = Pr\{X_n = i\}, \quad i = 1, \ldots, m$$

and n labels the time, and they are subject to the constraint $\sum_{i=1}^{m} \pi_n(i) = 1$. Let $\pi_n = (\pi_n(1), \pi_n(2), \ldots, \pi_n(m))$. These probabilities are calculated recursively via

$$\pi_{n+1} = \pi_n P$$

If the probabilities reach a steady state, then $\pi_{n+1} = \pi_n = \pi$, and we can simply solve the equation $\pi = \pi P$ subject to $\sum_{i=1}^{m} \pi(i) = 1$ to determine the steady state probabilities.

This is the most basic introduction to Markov processes; there is a great deal more information in the literature if you are interested.

Example: Social Mobility

Consider a town of 30,000 families which have been divided by the local chamber of commerce for planning purposes into three economic brackets (or classes): lower, middle, and upper. A builder wants to know the future populations of these brackets so they can decide what types (read: cost) of houses to build. The city also wants this information to help decide other issues like taxes and social services.

The past year the flow between these populations was the following:

- 20% of the lower move into the middle.
- 10% of the middle move back to the lower.
- 10% of the middle move to the upper.
- 15% of the upper move down to the middle.
- 5% of the lower move directly into the upper.
- 4% of the upper move down to the lower.

Last year there were 12,000 lower, 10,000 middle, and 8,000 upper income families.

Dynamical System Model

This situation can be modeled as a dynamical system. Let's use the following notation:

- x: number of families in upper income bracket,
- y: number of families in the lower income bracket,
- z: number of families in the lower income bracket,
- t: time measured in years.

We are assuming that the transitions between classes is a continuous event, and there is no immigration or emigration of the population. Also, since we are using a family as our basic unit of the population, we are not assuming that the families break apart or join. For short time periods this is reasonable, but for long time periods we we will be missing things like children starting their own families, deaths, births, and a host of other socioeconomic factors. But our model with these assumptions is a straightforward place to begin the analysis.

The state transition diagram is given in Fig. 1. The equations which govern this system are dynamical in nature, and we can assume that the flow is modeled by the initial value problem

$$\begin{aligned}
\frac{dx}{dt} &= -(0.15+0.04)x + 0.10y + 0.05z \\
\frac{dy}{dt} &= 0.15x - (0.10+0.10)y + 0.20z \\
\frac{dz}{dt} &= 0.05x + 0.10y - (0.05+0.20)z
\end{aligned}$$

where $x(0) = 8000$, $y(0) = 10000$, $z(0) = 12000$.

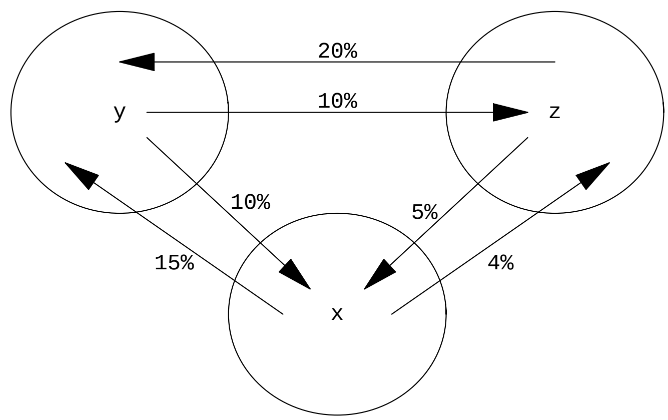

Figure 1: State transition diagram for the social mobility example. The self-to-self transitions have been left off the diagram.

The solution to this system of differential equations is

$$x(t) = \frac{4000\left(-364+233\sqrt{39}-\left(364+233\sqrt{39}\right)e^{\frac{\sqrt{39}\,t}{50}}+5850\,e^{\frac{(32+\sqrt{39})\,t}{100}}\right)}{2561\,e^{\frac{(32+\sqrt{39})\,t}{100}}}$$

$$y(t) = \frac{-10000\left(494+120\sqrt{39}+\left(494-120\sqrt{39}\right)e^{\frac{\sqrt{39}\,t}{50}}-3549\,e^{\frac{(32+\sqrt{39})\,t}{100}}\right)}{2561\,e^{\frac{(32+\sqrt{39})\,t}{100}}}$$

$$z(t) = \frac{4000\left(1599+67\sqrt{39}+\left(1599-67\sqrt{39}\right)e^{\frac{\sqrt{39}\,t}{50}}+4485\,e^{\frac{(32+\sqrt{39})\,t}{100}}\right)}{2561\,e^{\frac{(32+\sqrt{39})\,t}{100}}}$$

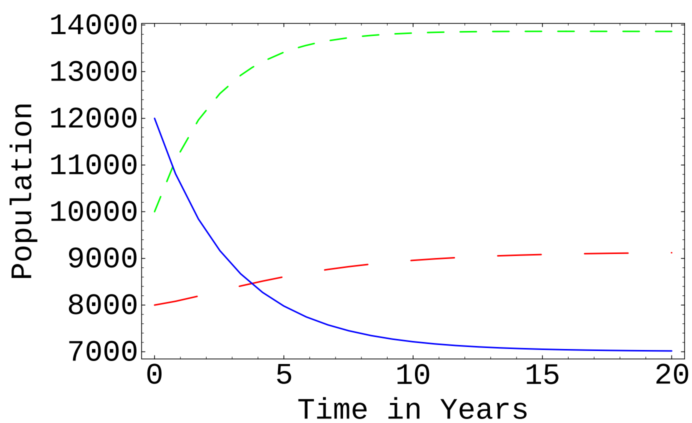

Figure 2: Dynamical system solution for the social mobility example. The red (large dash) line is the high economic class, the green (small dash) line is the middle economic class, and the blue (solid) line is the low economic class.

The equations are good to report, but certainly not very meaningful without further analysis. The results are also included in Fig. 2. From this figure, we see that the population in the middle class will increase in the coming years, and the population in the lower class will decrease. The population in the high economic class will increase, but its change is not as dramatic as the other two.

The final populations after 20 years are 9137 families in the high income, 13858 families in the middle income, and 7005 families in the low income. Of course, our model is probably not valid after 20 years, since the assumptions we have made will no longer be true.

For a more complete analysis of this dynamical system model, including sensitivity analysis, you should look at the *Mathematica* notebook at

`http://cda.mrs.umn.edu/~mcquarrb/Modeling/SocialMobility.nb`

Markov Process Model

The state transition diagram in Fig. 1 makes us think that we could model this as a Markov process. A Markov process is like a Markov chain, but events occur along a continuous time interval, and not a discrete time interval.

The probability matrix for the social mobility process is given by

$$P = \begin{pmatrix} 0.81 & 0.15 & 0.04 \\ 0.10 & 0.80 & 0.10 \\ 0.05 & 0.20 & 0.75 \end{pmatrix}$$

In this model, we could incorporate a flexible time process by assuming families move based on a Poisson process. However, that is bulky, and probably unnecessary. Instead, let's simply assume that we examine the population distributions once a year. This is a reasonable time step since the data is reported to us in terms of flow between the classes in the past year. We should not make our model more complicated unless we have to.

Our initial population distribution among the economic classes is given by:

$$\pi_0 = \left\{ \frac{8000}{30000}, \frac{10000}{30000}, \frac{12000}{30000} \right\} = \{0.267, 0.333, 0.400\}.$$

One year later, we can get the new population distribution by calculating

$$\pi_1 = \pi_0 P = \{0.269, 0.387, 0.344\}.$$

Now, we are really interested in populations, not distributions. So we should report the populations one year later, which are

$$30000 \times \pi_1 = \{8080, 11600, 10320\}.$$

This process can be continued, with $\pi_i = \pi_{i-1}P$. At each step the matrix P is *not* changed to reflect the new distribution of the population, since we are assuming the state transition diagram in Fig. 1 does not change. The populations are recorded in Table 1 and Fig. 3. Using this model we found (with the dynamical system model results are in parentheses) the final populations after 20 years are 9130 (9137) families in the high income, 13860 (13858) families in the middle income, and 7010 (7005) families in the low income. We see that we have remarkably good agreement with our dynamical system results.

Year	Population in High Economic Class	Population in Middle Economic Class	Population in Low Economic Class
0	8000	10000	12000
1	8080	11600	10320
2	8221	12556	9223
3	8376	13123	8502
4	8522	13455	8024
5	8649	13647	7704
6	8756	13756	7489
7	8842	13816	7342
8	8911	13847	7242
9	8965	13863	7173
10	9006	13869	7124
11	9038	13871	7090
12	9063	13871	7067
13	9081	13869	7049
14	9095	13868	7037
15	9106	13866	7029
16	9114	13864	7022
17	9120	13863	7018
18	9124	13862	7014
19	9127	13861	7012
10	9130	13860	7010

Table 1: Results of the Markov chain model for the social mobility problem.

Sensitivity Analysis of Markov Process

We need to perform some sensitivity analysis to determine how robust our solution is. For the social mobility problem, this can obviously take many forms. How much you do depends on the time you have to do it, and the importance of the results you are calculating.

Let's say that we want to make sure that our results are robust with respect to the initial state transition diagram we were given. We can perform many tests, but a simple one is to examine the steady state population transition probabilities when the initial probabilities are slightly changed.

Say we had an initial transition matrix given by:

$$P = \begin{pmatrix} 0.81 & 0.15 & 0.04 \\ 0.10 & 0.80 & 0.10 \\ 0.05 & (20+\lambda)/100 & (75-\lambda)/100 \end{pmatrix}$$

Mathematica can be used to solve for the steady state probabilities π. We find them to be:

$$\pi = \left\{ \frac{10\,(30+\lambda)}{985+29\,\lambda}, -\left(\frac{-455-19\,\lambda}{985+29\,\lambda}\right), \frac{230}{985+29\,\lambda} \right\}$$

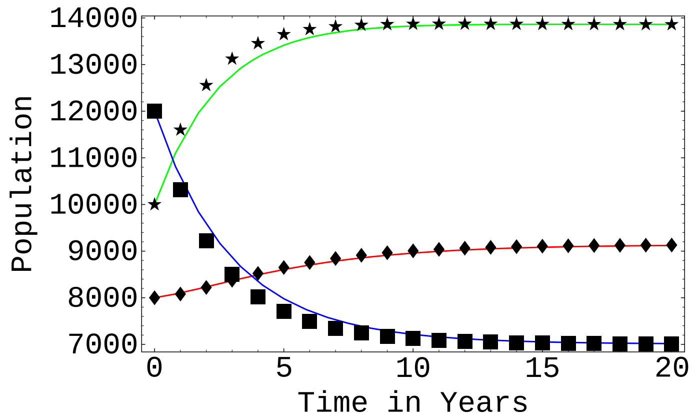

Figure 3: Markov chain solution for the social mobility example. The diamond points are the high economic class, the star points are the middle economic class, and the square points are the low economic class. The lines are the results from the differential equation solution to the same problem.

The important thing is to see if the steady state probabilities are sensitive to λ. Our analysis assumes $\lambda = 0$, and varying lambda by ± 10 would be a significant change in the initial state diagram. We therefore plot in Fig. 4 the steady state probabilities and see that they do not change drastically if we are close to zero. For $\lambda \sim -20$ the steady state probabilities do show rapid change, and other difficulties like a negative probability.

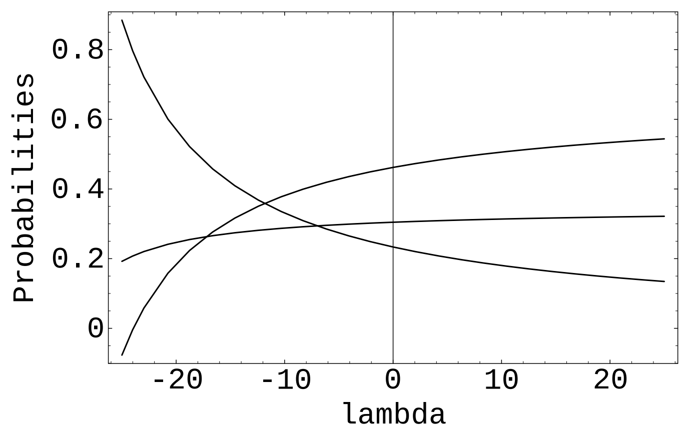

Figure 4: Sensitivity of the steady state probabilities to the parameter λ.

References

[1] Tommit Jaakkola, Machine learning and Neural Nets, MIT AI Lab, WebPage: http://www.ai.mit.edu/courses/6.867/lectures/lecture-15.pdf.

[2] H. Taylor & S. Karlin, An Introduction ot Stochastic Modeling, Academic Press (Orlando) 1984.

[3] M. Meerschaert, Mathematical Modelling, 2nd ed., Academic Press (San Diego) 1999.