ECE 6960: Adv. Random Processes & Applications
Lecture Notes, Fall 2010

Lecture 20

Today: (1) Project (2) Continuous Time Markov Chains

- HW 9 due Thursday.
- I was asked to attend an NSF grantees conference Nov. 30 through Dec. 3. This will mean that I need to miss two lectures, on Nov 30 and Dec 2. I have eliminated one topic (Bayesian networks) and condensed three lectures into two. This way, I can avoid having makeup lectures at a time when you should be working on your projects. I would be happy to explain Bayesian networks to you if you are interested in that topic, come and see me in office hours.

1 Project

From the syllabus:

- Students will do an individual course project, including a written report, and a oral presentation. The oral presentations will be presented in class during the final two class meetings. The written report will be due December 9, the final day of class. The course project requires students to apply the tools of the course to a particular application. A project may involve either experiment or simulation, and must apply the tools from the course to analyze or model a random process. The project must involve a random process, rather than simply a random variable. From the beginning of the semester, students should scan the readings in an effort to understand the breadth of the course topics in order to formulate ideas for an interesting and challenging project.

Here are some further rules.

1.1 Written Report

For the written project:

- Cite all sources, and cite properly. ALL SUBMITTED REPORTS ARE AUTOMATICALLY CHECKED BY PLAGIARISM DETECTION SITE TURNITIN.COM. You may use whatever sources you wish, but you MUST properly cite the source of your ideas. When you paraphrase someone else's words from a journal paper or book, cite that source after your paraphrased explanation. In engineering, 99% of citations are a paraphrase citation. When you need to copy someone's words exactly, put their words in quotes and cite the source after the end of the quotation. Don't copy more than a sentence. If you do really need to copy an entire paragraph from a source, make it an indented quote block. I have never seen a case when a student really needed to copy an entire paragraph from a source. I've only seen cases when a student was using this means to avoid writing their own words. I don't give you "credit" for copying someone else's words in terms of the grade for your written project. If a third of your paper is copied quotes, then you essentially have neglected to write a third of your written report.

- Related to the previous note: HW 10 will be a 20 minute "plagiarism assignment" to be completed after Exam 2 (see the link on the schedule site).
- Use a minimum of three references. If I don't set a minimum, some people will not cite any sources. This is a experimental or simulation report. You are getting some ideas for the project from somewhere, so cite these sources. You can cite the book, or my class notes, or some journal or conference papers, or a web site, even. If you are solely recreating the work of one paper, please look for other papers which cite this paper (use Google Scholar) and see what they have done to expand on its results. Or look at the citations of that paper and describe what that paper's contribution was. Again, paraphrase and then cite the work.
- For your report, I would recommend sections:
    1. Introduction
    2. Methods
    3. Results
    4. Discussion
    5. Conclusion
- Consider the "Methods" section. You should be able to describe in detail the tools and procedures that you used to do your simulation or experiment or analysis. There should be enough detail that I could recreate your work exactly. So for example, if you are are simulating a Markov chain, then you should provide enough detail to see how you set every parameter in that MC, including initial probabilities, transition

probabilities, etc. You should describe an overview of your code as well, at a high level. You may include Matlab code as an appendix, but don't put your code in the main body of the report. If you are going to test for independence, or compare an experimental CDF to an analytical CDF, describe the math in this section. You will refer to this math in the Results and Discussion sections.

- In the results section, you might put output from the simulation, or results from processing your experimental data. Put in figures, but describe the figures in the caption or text, saying what parameters were used to generate each figure.
- Do not present results without any discussion. In the discussion section, describe on a higher level what was achieved, and what you learned. For example, if the body of results shows that as you increase parameter A, that performance measure B goes down as well, make that argument here. Describe why that does or doesn't make sense, given your knowledge of the field. You might have two or three things that one can learn from the results. If you want to present results and discussion together in one section, that is fine as well.
- I don't like to give you a minimum number of pages, but every student wants one. I will say that six pages should be sufficient (excluding any code included in an appendix) for most projects, although some will naturally require more pages to fully explain the project.
- The written report is due midnight on Dec. 9.

If you want more technical writing tips (opinions) I have put up "Twenty-eight tips for research writing" on the web at: `http://span.ece.utah.edu/tips_research_writing`.

1.2 Presentation

For your presentation, you will present slides.

1. Presentations will be done December 7 and 9, during class time.
2. We have eleven student taking this class for a grade, and 160 minutes in total. So, plan for a ten minute presentation with a couple of minutes for questions, and then a couple minutes to switch laptops or presenter.
3. You will be grading other students' presentations, so you must attend both class sessions.

4. You should build on what students in this class already know about random processes. Your project is related to one or more topics presented in class, so emphasize those connections.

5. The presentation is at least a little bit before the written report is due, so be sure to be done with your project work by your presentation date.

1.3 Grade

The total grade will be assigned as 1/3 from the presentation, and 2/3 from the written report. The project overall is 30% of your final grade. Don't be afraid to chose a more challenging project. If you do a less challenging project, I will expect it to be more complete and more accurate. If you chose a more challenging project, I will be more forgiving of you not explaining every aspect of each result; and more forgiving of a mistake.

Please do contact me with your project topic and plan, and consult with me as you obtain results. I am happy to help you come up with things to do or try next, or to help you understand your results. Please do not wait until the last day.

2 Continuous-time Markov Chains

This is Section 11.4 in Leon-Garcia. We're going to discuss continuous time Markov chains, that is, a cts-time R.P. $X(t)$ with the Markov property that is discrete-valued, and stationary. When $X(t)$ is Markov, we need to find a model for

$$P\left[X(s+t)=j|X(s)=i\right]$$

for $t \geq 0$. Because of the assumed stationarity, the model is only a function of the time difference between the two samples, in this case, t. Thus we use the shorthand notation,

$$p_{i,j}(t) = P\left[X(s+t)=j|X(s)=i\right]$$

For any given t, we can form a matrix of transition probabilities, $P(t)$, which the (i,j)th element of $P(t)$ describes the probability of transitioning from state i to state j after time t. Note $P(0) = \mathrm{I}$ since $P\left[X(s)=j|X(s)=i\right]=0$ for $i \neq j$ and $P\left[X(s)=i|X(s)=i\right]=1$.

Note a Poisson process can be described as a cts-time Markov chain. Same goes for the random telegraph wave process. For these two processes, it is easy to see that the time spent in each state is exponential, since interarrivals in the Poisson R.P. are exponential.

2.1 Time Spent in a State

However, we also know that the *time spent* in any given state i in a cts-time Markov chain also has the exponential distribution. The time spent in a state is the time duration between arriving in state i and leaving state i. This is essentially because the Markov property implies that it doesn't matter how long the MC has been state i, that the conditional probability of being in state i at a time t in the future doesn't change by knowing how long $X(s)$ has already been in state i. In other words, for $s_0 < s_1$ and $t \geq 0$,

$$P\left[X(s_1+t) = j | X(s_1) = i, X((s_0, s_1) = i\right] = P\left[X(s_1+t) = j | X(s_1) = i\right]$$

Only the exponential pdf has this "memoryless" property. Describing the time spent in state i as T_i, then,

$$P\left[T_i > t\right] = e^{-v_i t}$$

for some *mean state occupancy time* $1/v_i$.

The above result is used to provide a second perspective on a cts-time Markov chain. This perspective says that each time a state i is entered, we can view the process as "selecting" an exponential r.v. T_i to describe the duration in that state. Then, a regular (discrete-time) MC is used to select the next state, with transition probability $\tilde{q}_{i,j}$ for each other state j, independent of T_i. This discrete-time MC with TPM $P(1) = \tilde{Q}$ is called the "embedded Markov chain". This is a good perspective on how to simulate a cts-time Markov chain.

2.2 Rates of Arrival and Departure

For cts-time Markov chains, we typically describe them in terms of the rate of transition from one state to another. Questions about rate are questions about derivatives. Since we know that $P\left[T_i > t\right] = e^{-v_i t}$, what is the rate of departing from state i? The slope of $P\left[T_i > t\right]$ at time zero is $-v_i$. We call v_i the rate of departure. In another perspective, consider the probability of remaining in state i after a very small duration δ:

$$p_{i,i}(\delta) = P\left[T_i > t\right] = e^{-v_i t} \approx 1 - \frac{v_i \delta}{1!} + \frac{v_i^2 \delta^2}{2!}$$

Thus neglecting higher order terms,

$$p_{i,i}(\delta) \approx 1 - v_i \delta$$

So for very small δ, the MC leaves state i with rate v_i. Once the process leaves state i, it enters other states with transition probabilities $\tilde{q}_{i,j}$ independent of the time of departure.

$$p_{i,j}(\delta) = (1 - p_{i,i}(\delta))\tilde{q}_{i,j} \approx v_i \tilde{q}_{i,j} \delta$$

We define a new variable $\gamma_{i,j} = v_i \tilde{q}_{i,j}$ and describe it as the rate at which the process $X(t)$ enters state j from state i. To be complete, we define $\gamma_{i,i} = -v_i$ because v_i is the rate at which the process $X(t)$ leaves state i.

2.3 State Probability Vector and Chapman-Kolmogorov

Let's also define the state probability vector as we did for discrete-time Markov chains. Now, let $p_j(t)$ be the probability of being in state j at time t,

$$p_j(t) = P\left[X(t) = j\right]$$

Side note: Yes, the notation is confusing with the pervasive use of the letter "p". If we talk about uppercase $P(t)$, we are talking about a transition probability matrix. If we have a lowercase p, it depends on how many subscripts it has. If it has one subscript, *i.e.*, $p_j(t)$, we are talking about the probability of being in state j at time t. If it has two subscripts, like $p_{i,j}(t)$, we are talking about one element of the transition probability matrix, the probability of transitioning from state i to state j in t time units.

That being said, we can re-derive a version of the Chapman-Kolmogorov equations for cts-time Markov chains:

$$\frac{\partial}{\partial t} p_j(t) = \sum_i \gamma_{i,j} p_i(t)$$

Note that solving these equations for all j requires solving a set of differential equations, subject to the initial conditions given by $p_j(0)$, for all j. The Leon-Garcia book has two good examples of deriving the state probability vector as a function of t in section 11.4.

2.4 Stationary pmf

Finally, as $t \to \infty$, systems that reach "equilibrium" converge to a pmf, just like a discrete-time MC has a stationary probability vector $\boldsymbol{\pi}$. Of course, not all cts-time Markov chains will converge to a unique equilibrium state. But for those which do, the rate of change goes to zero. Denoting $p_i = \lim_{t\to\infty} p_i(t)$, when it exists,

$$0 = \sum_i \gamma_{i,j} p_i$$

because we've specified that $\gamma_{i,i} = -v_i$, then

$$v_j p_j = \sum_{i \neq j} \gamma_{i,j} p_i$$

This is essentially a "sum of the inputs equals the sum of the outputs" equation like you've seen in your basic circuits classes.