ESE303 Midterm II Study Guide

Kate Tolstaya, Vinicius Lima*

December 9, 2019

1 Continuous Time Markov Chains

1.1 Memoryless property: continuous time

Markov chains are characterized by the Markov or memoryless property:

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

This is a discrete-time formulation, how do we extend it to a continuous time setting?

$$P(T > s+t | T > s) = P(T > t)$$

Note that we can also write

$$P(T > s+t | T > s) = \frac{P(T > s+t, T > s)}{P(T > s)} = \frac{P(T > s+t)}{P(T > s)}$$

which implies that

$$P(T > s+t) = P(T > t)P(T > s).$$

This way it is easier to see how the memoryless property is related to exponential random variables, that we discuss next. Recall that memoryless random times need to have exponential distribution.

1.2 Exponential RVs and Poisson Processes

Exponential RV - time elapsed between occurrence of random events, continuous RV $T \sim exp(\lambda)$ is exponential with parameter $\frac{1}{\lambda}$. Its PDF is:

$$f_T(t) = \lambda e^{-\lambda t}$$

Its expectation is $\frac{1}{\lambda}$, the average rate of events happening at intervals T. The variance is $\frac{1}{\lambda^2}$
Its cumulative distribution function (CDF), on the other hand, is given by

$$F_T(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

Note that the complement of the CDF measures the probability of the random variable T being greater than some value t:

$$P(T > t) = e^{-\lambda t}$$

Moreover we can see that exponential RVs are memoryless since

$$P(T > s+t) = e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t} = P(T > t)P(T > s)$$

If $T = \min(T_1, T_2)$, the time to the first event, where $T_1 \sim exp(\lambda_1)$ and $T_2 \sim exp(\lambda_2)$, then

$$P[T > t] = P[T_1 > t]P[T_2 > t] = e^{-(\lambda_1 + \lambda_2)t}$$

A Poisson RV measures the number of occurrences of random events in a fixed time interval with rate (parameter) λ. A poisson RV has PMF:

$$P[N=k] = \frac{e^{-\lambda}(\lambda)^k}{k!}$$

and mean λ.
Poisson Process - number of occurrences of independent events in a time interval of length T with rate (parameter) λ

$$P[N(T)=k] = \frac{e^{-\lambda T}(\lambda T)^k}{k!}$$

The inter-arrival times of the Poisson Process with parameter λ are independent exponential RVs with parameter λ.
Recall that a Poisson Process has independent (that is the number of events in any pair of disjoint intervals are independent) and stationary (probability distribution of the number of events depends on the length of the interval but not on its absolute position) increments. It is an example of a simple CTMC.

1.3 CTMC

Continuous time: nonnegative random variable $t \in \mathbb{R}^+$, States $X(t)$ take countable values $0, 1, ...$
Continuous time, countable state stochastic process X(t) is a CTMC if the memoryless property holds (the future is independent of the past given the present)

$$P[X(t+s) = j | X(s) = i, X(u) = x(u), u < s] = P[X(t+s) = j | X(s) = i]$$

Applying the memoryless property to the transition times T_i, time until transition out of state i into any other state j

$$P[T_i > t+s | T_i > s] = P[T_i > t]$$

Transition times are exponentially distributed $T_i \sim exp(\nu_i)$. ν_i is the rate of transitioning out of state i. When a transition happens, X(t) cannot return to state i (no self-loops)

Here we focus on homogeneous MCs, in which the transition probabilities do not depend on the starting point s but only on the length of the time interval t

2 ways to specify CTMC:
1) Transition rates out of each state ν_i, Transitions probabilities out of state i into state j, P_{ij}:

$$\sum_{j=1}^{\infty} P_{ij} = 1, P_{ii} = 0$$

2) Transition rates out of state i to state j, $q_{ij} = \nu_i P_{ij}$
To convert back to 1:

$$\nu_i = \sum_{j=1, j\neq i}^{\infty} q_{ij}, P_{ij} = q_{ij}/\nu_i$$

Note that so far we discussed a sort of algorithmic definition of CTMCs: transition times from state i are exponential, and when they occur, they happen into state j with probability P_{ij}. Now we wanna find the transition probability function $P_{ij}(t)$, that can be found by using the Chapman-Kolmogorov equations.

1.4 Chapman-Kolmogorov Equations

Our goal is to find $P_{ij}(t)$, the probability of being at state j at time t, if $X(0) = i$
The Forward (where the process is going to) and Backward (where the process is coming from) equations are derived from theorem:

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s)$$

Note that here we are considering the probability of going from state i to k at time t, then going from state k to j in the remaining time s and summing over all possible intermediate steps k.
Forward equations:

$$\frac{\partial P_{ij}(t)}{\partial t} = \sum_{k=0,k\neq j}^{\infty} q_{kj}P_{ik}(t) - v_j P_{ij}(t)$$

Left-hand side = rate of change
$q_{kj}P_{ik}(t)$ = (transition into k in $0 \to t$) $\times$(rate of moving into j in next instant)
$v_j P_{ij}(t)$ = (transition into j in $0 \to t$) $\times$ (rate of leaving j in next instant)
Change in $P_{ij}(t) = \sum_k$(moving into j from k) $-$ (leaving$j$)

Backward equations:

$$\frac{\partial P_{ij}(t)}{\partial t} = \sum_{k=0,k\neq i}^{\infty} q_{ik}P_{kj}(t) - v_i P_{ij}(t)$$

Change in $P_{ij}(t) = \sum_k$((moving into k from i) $-$ (rate from i to j)

Matrix formulation of the 2 equations:
Forward equation: P'(t) = P(t)R, Backward equation: P'(t) = R P(t)
where P(t) = $\begin{bmatrix} P_{11}(t) & \dots & P_{1N}(t) \\ \vdots & \ddots & \vdots \\ P_{N1}(t) & \dots & P_{NN}(t) \end{bmatrix}$ and the elements of R are given by $R_{ij} = q_{ij}$ and $R_{ii} = -v_i$
Solve for $P(t) = e^{Rt}$, the matrix exponential of Rt

Given $P(t)$ and initial distribution $p(0)$, we can find unconditional probabilities as

$$p(t) = P^T(t)p(0). \tag{1}$$

Now recall that we can define an embedded (discrete-time) MC w/ transition probabilities P with null diagonal:
CTMC irreducible if MC single class
transient and recurrent states in CTMC if transient/recurrent in MC

(Theorem) For a irreducible positive recurrent CTMC with transition rates v_i and $q_{ij}, P_j = \lim_{t\to\infty} P_{ij}(t)$ exists for all i, is independent of initial state i and satisfies

$$v_j P_j = \sum_{k=0,k\neq j}^{\infty} q_{kj}P_k, \ \sum_{j=0}^{\infty} P_j = 1 \tag{2}$$

If CTMC is irreducible, positive recurrent, then ergodicity holds; so ergodic limit coincides with limit probabilities
To find the limit probabilities P_i, the probability of X(t) being in state i in steady state, where $\frac{\partial P_{ij}(t)}{\partial t} = 0$, so we can simplify the Kolmogorov equations — Balance equations:
For each state j at steady state, rate into j = rate out of state j, $\sum_i q_{ij}P_i = \sum_k q_{jk}P_j$

2 Queues

2.1 M/M/1 Queue

Markov Arrivals, Markov Departures, 1 Server. Customers arrive at rate λ and leave with rate μ (independent)
We transition from state with i customers whenever an arrival or departure happens (for $i \neq 0$): $v = \lambda + \mu$
Arrival happens before departure with probability $P_{i,i+1} = \frac{\lambda}{\lambda+\mu}$
Departure happens before arrival with probability $P_{i,i-1} = \frac{\mu}{\lambda+\mu}$
For $i = 0$. Special case - customers can only arrive, can't leave: $v_0 = \lambda, P_{01} = 1$

2.2 Other Queues

Multiserver queues: M/M/c; Poisson servers with independent service rates
Networks of Queues: customers arrive at system to receive two services. Service 1 is performed with rate μ_1 and service 2 with rate μ_2. Behaves as two independent M/M/1 queues.

3 Biochemical Reactions

The system's sate is $X(t)$, which tells the amount of each molecule. To specify the i-th reaction, quantity of reactants, reaction rate, quantity of products

$$R_i : s_{i1}^l X_1 + \ldots + s_{im}^l X_m \xrightarrow{c_i} s_{i1}^r X_1 + \ldots + s_{im}^r X_m$$

To write n reaction equations as a system of equations, $S^{(l)}X \xrightarrow{c_i} S^{(r)}X$

Left stoichiometry matrix: $S^{(l)} = \begin{bmatrix} s_{11}^l & \cdots & s_{1m}^l \\ \vdots & \ddots & \vdots \\ s_{n1}^l & \cdots & s_{nm}^l \end{bmatrix}$ Right stoichiometry matrix: $S^{(r)} = \begin{bmatrix} s_{11}^r & \cdots & s_{1m}^r \\ \vdots & \ddots & \vdots \\ s_{n1}^r & \cdots & s_{nm}^r \end{bmatrix}$

Hazards $h_i(X)$ determine the transition rates of the underlying CTMC.

- Hazards for zero-th, first and second order reactions (for reference)

Order	Reaction	Rate
zero-th	$\emptyset \xrightarrow{c} \cdots$	c
first	$X_1 \xrightarrow{c} \cdots$	cX_1
second	$X_1 + X_2 \xrightarrow{c} \cdots$	cX_1X_2
second	$2X_1 \xrightarrow{c} \cdots$	$cX_1(X_1-1)/2$

Gillespie's algorithm: simulation of CTMC / biochemical reactions. Calculates hazards. Calculate transition rate. Random (exponential) time of next reaction. Advances time by δ_t. Draw reaction ($\frac{h_i(X)}{v(X)}$). Update state vector.

4 Gaussian processes

Stochastic processes: Markov (memoryless) vs Gaussian (normal distribution) vs Stationary (probabilities are invariant to time shifts) $X(t)$ is a Gaussian process when all prob. distributions are Gaussian

for any $t_1,\ldots,t_n \to X(t_1),X(t_2),\ldots,X(t_n)$ are jointly Gaussian
RVs $X_1,\ldots,X_n$ are jointly Gaussian if any linear combination of them is Gaussian
More general, any linear functional of X(t) is normally distributed

Recall that linear transformation of independent Gaussians is Gaussian.

To specify a **Gaussian process**:
mean value function : $\mu(t) = \mathbb{E}[X(t)]$
autocorrelation function: $R(t_1,t_2) = \mathbb{E}[X(t_1)X(t_2)]$; symmetric: $R(t_1,t_2) = R(t_2,t_1)$
if needed, autocovariance: $C(t_1,t_2) = R(t_1,t_2) - \mu(t_1)\mu(t_2)$

4.1 Brownian Motion

1. $X(t) \sim \mathcal{N}(0,\sigma^2 t)$
2. Independent increments: $X(t_2) - X(t_1)$ ind. of $X(s_2) - X(s_1)$ for disjoint (t_1,t_2) and (s_1,s_2)
3. stationary increments: $X(t+s) - X(s)$ and $X(t)$ have same distribution
4. autocorrelation: $R_X(t_1,t_2) = \sigma^2 min(t_1,t_2)$

4.2 White Gaussian noise (WGN) process $W(t)$

1. example of a gaussian process
2. zero mean, $\mathbb{E}[W(t)] = 0,\ \forall t$
3. autocorrelation depends on the delta function: $R_W(t_1,t_2) = \sigma^2\delta(t_1 - t_2)$
4. (delta function $\delta(t) = \begin{cases}\infty, & t = 0\\ 0, & \text{else}\end{cases}$
5. $\int_a^b f(t)\delta(t)dt = \begin{cases} f(0), \text{ if } a < 0 < b \\ 0 \text{ o.w.}\end{cases}$)
6. here $R_W(t_1,t_2) = \mathbb{E}[W(t_1)W(t_2)] = 0$ if $t_1 \neq t_2$
7. Values of $W(t)$ at different times are independent! Since $W(t)$ is Gaussian, uncorrelation implies independence:
   (a) Independent $\Rightarrow$ uncorrelated
   (b) In general, Uncorrelated $\not\Rightarrow$ independent
   (c) for **Gaussian** Processes, uncorrelated $\Leftrightarrow$ independent

4.3 Integral of WGN

Consider a GP

$$X(t) = \int_0^t W(u)du$$

Integration is a linear functional $\implies$ X(t) is also a GP!
Mean function $\mu(t)$:

$$\mu(t) = \mathbb{E}[X(t)] = \mathbb{E}[\int_0^t W(u)du] = \int_0^t \mathbb{E}[W(u)]du = 0$$

Compute the autocorrelation function $R(t_1,t_2)$ of the same GP

$$\begin{aligned}
R_X(t_1,t_2) &= \mathbb{E}\left[\left(\int_0^{t_1} W(u_1)du_1\right)\left(\int_0^{t_2} W(u_2)du_2\right)\right] \\
&= \mathbb{E}\left[\int_0^{t_1}\int_0^{t_2} W(u_1)W(u_2)du_1du_2\right] \\
&= \int_0^{t_1}\int_0^{t_2} \mathbb{E}[W(u_1)W(u_2)]du_1du_2 \\
&= \int_0^{t_1}\int_0^{t_2} \sigma^2\delta(u_1 - u_2)du_1du_2 \ \ (\text{assume } u_1 < u_2) \\
&= \int_0^{t_1} \sigma^2 du_1 = \sigma^2 t_1 \ \ (\text{similar when } u_2 < u_1).
\end{aligned}$$

$\implies$ same mean and autocorrelation as BM!

5 Pricing of options and stocks

Arbitrage = " It is possible to devise a betting strategy that guarantees a positive return no matter the combined outcome of the events"
Parameters / variables:

1. $k = 1,\ldots,K$: events
2. $j = 1,2,\ldots,J$: joint outcomes ("world realization")

3. r_{jk}: return (per unit)
4. x_k: bet in outcome k
5. Total return = $\mathbf{x}^T \mathbf{r}_j$

Arbitrage theorem:

Theorem
Given vectors of returns $\mathbf{r}_j$, *associated with random outcome* $j = 1, \ldots, J$ *an arbitrage is not possible if and only if there exist a probability vector* $\mathbf{p}$ *such that* $\mathbb{E}_{\mathbf{p}}(\mathbf{r}) = \mathbf{0}$. *Equivalently,*

$$\max_{\mathbf{x}} \left(\min_j \left(\mathbf{x}^T \mathbf{r}_j \right) \right) \leq 0 \quad \Leftrightarrow \quad \sum_{j=1}^{J} p_j \mathbf{r}_j = \mathbf{0}$$

- Prob. vector $\mathbf{p}$ is **NOT** the prob. distribution of events $j = 1, \ldots, J$

5.1 Returns

Expected non-discounted return r of an investment in a stock

1. $X(t)$ be the price of a stock, where we assume follow a geometric Brownian motion with drift μ and volatility σ^2
2. Can define a $Y(t)$ Brownian motion with drift μ and variance σ^2 such that $X(t) = X(0)e^{Y(t)}$
3. Expected non-discounted return r of an investment in a stock is $\mathbb{E}[e^{Y(t)}] - 1$
4. $\mathbb{E}[e^Y] = e^{\mu_Y + \sigma_Y^2/2}$

Continuously compounded return α by investing in the money market (risk free)

1. Invest one dollar, gives you $e^{\alpha t}$ at time t

Discounted return of an investment in a stock

1. Return discounted compared to the money market rates, continuously compounded return α by investing in the money market
2. If non-discounted return $\frac{\mathbb{E}[X(t)]}{X(0)} - 1$, discounted return is $e^{-\alpha t} \frac{\mathbb{E}[X(t)]}{X(0)} - 1$

5.2 Risk neutral measure q

q is a joint probability distribution. The expected value of earnings w.r.t **q** is zero if no arbitrage!

1. We consider the change over 1 period of h:
2. $X((n+1)h) = X(nh)e^{\sigma\sqrt{h}Y_n}$
3. Each element Y_n is a binary random variable with probabilities

$$P(Y_n = 1) = q, \quad P[Y_n = -1] = 1 - q$$

4. Examine the change during the first interval h:

$$P[X(h) = X(0)e^{\sigma\sqrt{h}}] = q, \quad P[X(h) = X(0)e^{-\sigma\sqrt{h}}] = 1 - q$$

5. Consider an investment $X(0) = \$1$,

$$r_1 = e^{\sigma\sqrt{h}} - 1, \quad r_2 = e^{-\sigma\sqrt{h}} - 1$$

6. The only situation where there is no arbitrage is when

$$qr_1 + (1-q)r_2 = 0$$

7. Solving for q yields

$$q = \frac{e^{\alpha h} - e^{-\sigma\sqrt{h}}}{e^{\sigma h} - e^{-\sigma\sqrt{h}}}$$

8. For small h, $e^{\alpha h} \approx 1 + \alpha h$ and $e^{\pm\sigma\sqrt{h}} \approx 1 \pm \sigma\sqrt{h} + \sigma^2 h/2$, hence

$$q \approx \frac{1}{2}\left(1 + \frac{\alpha - \sigma^2/2}{\sigma}\sqrt{h}\right)$$

9. Thus, measure q $:= \lim_{h\to 0} q(h)$ is geometric Brownian motion with variance σ^2 (same as stock price) and drift $\alpha - \sigma^2/2$

10. Expected growth:

- Compute expected return on an investment on stock $X(t)$
- Buy 1 share of stock at time 0. Cash invested $\Rightarrow X(0)$
- Sell stock at time t. Cash after sell $\Rightarrow X(t)$
- Expected value of cash after sell given $X(0)$ is

$$\mathbb{E}\left[X(t) \,\middle|\, X(0)\right] = X(0)e^{(\mu+\sigma^2/2)t}$$

- Alternatively, invest $X(0)$ risk free in the money market
- Guaranteed cash at time t is $X(0)e^{\alpha t}$
- Invest in stock only if $\mu + \sigma^2/2 > \alpha$ $\Rightarrow$ risk premium

5.3 Arbitrage

The no-arbitrage price c of the option on the stock $X(t)$ with strike time t and strike price K as an expected value w.r.t to the risk neutral measure

1. Option is a contract to buy shares of a stock are a future time
2. May become active or not, depending if $(X(t) - K) > 0$
3. Option values at time t, $(X(t) - K)^+$
4. Return value at time 0, $e^{-\alpha t}(X(t) - K)^+ - c$
5. To have no arbitrage, the option's price should be

$$\mathbb{E}_{\mathbf{q}}\left[e^{-\alpha t}(X(t) - K)^+ - c\right] = 0$$

6. In terms of c,

$$c = e^{-\alpha t}\mathbb{E}_{\mathbf{q}}\left[(X(t) - K)^+\right]$$

6 Linear filtering

6.1 Stationary processes

$\approx$ study of limit distributions

1. All probabilities are invariant to time shifts: strictly stationary

$$P[X(t_1+s)\geq x_1,\ldots,X(t_n+s)\geq x_n]=P[X(t_1)\geq x_1,\ldots,X(t_n)\geq x_n]$$

2. A much weaker condition first order stationary $\implies$ probabilities of single variables are shift invariant

$$P[X(t+s)\geq x]=P[X(t)\geq x]$$

3. For SS processes, pdfs and joint cdfs are shift invariant
   $\implies$ mean, variance, power of a SS process is constant
   $\implies$ autocorrelation depends on t_2-t_1 only: $R_X(t_1,t_2)=R_X(0,t_2-t_1)=R_X(s)$
4. For GP, implying $\mu(t+s)=\mu(t)$ and $R_X(t,t)=R_X(t+s,t+s)$

6.2 Wide sense stationary (WSS)

1. WSS: not SS, mean is constant, autocorrelation shift invariant $R_X(t_1,t_2)=R_X(t_1-s,t_2-s)$
2. Power $\mathbb{E}[X^2(t)]$ of a process

$$\mathbb{E}[X^2(t)]=R_X(0)=R_X(t,t)=\mathbb{E}[X(t)X(t)],\text{ for WSS process}$$

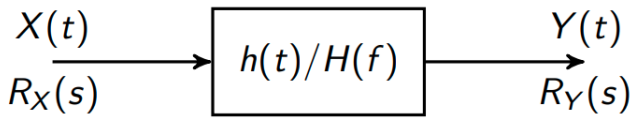
3. For Gaussian Processes, WSS and SS are equivalent

6.3 Power spectral density

The power spectral density (PSD) of a stochastic process is the Fourier transform of the autocorrelation function

1. $S_X(f)=\mathscr{F}(R_X(s))$
2. e.g. $\mathscr{F}(\cos(2\pi f_0 x))=\frac{1}{2}\delta(f-f_0)+\frac{1}{2}\delta(f+f_0)$

Input signal $X(t)$ passes through a linear filter H



PSD relation:

- Note definitions of $\Rightarrow X(t)$'s PSD $\Rightarrow S_X(f)=\mathcal{F}(R_X(s))$
  $\Rightarrow$ Filter's frequency response $\Rightarrow H(f):=\mathcal{F}(h(t))$
  Also note that $\Rightarrow H^*(f):=\mathcal{F}(h(-t)))$
- Latter three observations yield (also use $H(f)H^*(f)=|H(f)|^2$)

$$S_Y(f)=H(f)S_X(f)H^*(f)=|H(f)|^2S_X(f)$$

For

$$R_X(t_1,t_2)=\cos(2\pi f_0(t_2-t_1))+\cos(8\pi f_0(t_2-t_1))$$

with cut out frequency of H at $3f_0$, we know that

$$R_Y(t_1,t_2) = \cos(2\pi f_0(t_2 - t_1))$$

since the high frequency is cut out.