Continuous-time Markov Chains

6.2 Introduction

- **continuous-time stochastic process** $\{X(t), t \geq 0\}$
  - state: finite/countable values, $\{0{,}1{,}2, \ldots\}$
  - time: continuous, $t \geq 0$
  - transition prob (time stationary/homogeneous), $\Pr(X(t+s) = j | X(s) = i, X(u), 0 \leq u < s) = \Pr(X(t) = j | X(0) = i)$.
  - T_i denote the amount of time that the process stays in state i before making a transition into a different state
    - $\Pr(T_i > s + t | T_i > s) = \Pr(T_i > t)$
      - the process is in state i at time s
      - by the Markovian property, that the prob that it remains in that state during the interval [s,s+t] is just the (unconditional) prob that it stays in state i for at least t time.
    - T_i is memoryless
- **memoryless continuous random variable X**: $\Pr(X > t + s | X > s) = \Pr(X > t)$ <==> $\Pr(X > t + s) = \Pr(X > t)\Pr(X > s)$
  - survival function of X: $G(t) = \Pr(X > t)$ (right continuous)
  - G(t+s)=G(t)G(s)
    - easy to check $G(nt) = G(t)^n$, or $G(t) = G(nt)^{1/n}$.
    - so $G(1/p) = G(1)^{1/p}$, $G(q/p) = G(1/p)^q = G(1)^{q/p}$ for any integer p,q.
    - rational number can approximate any real number, so $G(t) = G(1)^t = \exp(-\log(1/G(1))t)$, an exponential dist with rate $\lambda = \log(1/G(1))$.
  - *geometric dist* is the discrete analog: survival prob $(1-p)^k$.
  - *exponential dist with rate* λ: $f(t) = \lambda e^{-\lambda t}$, $G(t) = e^{-\lambda t}$; mean $E[T] = 1/\lambda$, variance $Var[T] = 1/\lambda^2$.
  - minimum of independent exp is still exp (rate is the sum of individual rates).

                - $\Pr(\min(X, Y) > t) = \Pr(X > t)\Pr(Y > t) = e^{-t\lambda_x - t\lambda_y}$
                    - $\Pr(\min_i Y_i > t) = \prod_i \Pr(Y_i > t) = e^{-t\sum_i \lambda_i^y}$
                - Pr(an exp is the minimum among a set of independent exps) $\propto$ its rate
                    - $\Pr(X < Y) = E[E[I(X < Y)|X]] = E[e^{-X\lambda_y}] = \lambda_x/(\lambda_x + \lambda_y)$
                    - $\Pr(X < Y_1, \cdots, X < Y_m) = E[E[I(X < Y_i, \forall i)|X]]$
                        - $= E[e^{-X\sum_i \lambda_i^y}] = \lambda_x/(\lambda_x + \sum_i \lambda_i^y)$
        - sum of iid exp is gamma(n,λ).
        - Failure/hazard rate
            - r(t)=f(t)/G(t)
            - $\Pr(X \in (t, t + dt)|X > t) = r(t)dt$.
            - $G(t) = \exp(-\int_0^t r(t)dt)$.
            - $r(t) = \lambda$ for exponential dist.

- **alternative definition of continuous-time MC**
    - a stochastic process that moves from state to state in accordance with a (discrete-time) MC
        - $P_{ii} = 0, \sum_{j \neq i} P_{ij} = 1$
    - the amount of time it spends in state i, before proceeding to the next state, is exponentially distributed with rate v_i
    - the amount of time the process spends in state i, and the next state visited, must be independent random variables

---

6.3 Birth and Death Processes

- **Birth and death process**
    - a system whose state at any time is represented by the number of people in the system at that time
    - whenever there are n people in the system
        - new arrivals enter the system at an exponential rate λ_n (*birth/arrival rate*)
        - people leave the system at an exponential rate μ_n (*departure/death rate*)
            - time until the next arrival is exponentially distributed with mean $1/\lambda_n$,
            - independent of the time until the next departure, which is itself exponentially distributed with mean $1/\mu_n$.

    - a continuous-time MC with states {0,1,…} for which transitions from state n may go only to either state n-1 or state n+1
        - $\nu_0 = \lambda_0, \nu_i = \lambda_i + \mu_i, i > 0.$
        - $P_{01} = 1, P_{i,i+1} = \lambda_i/(\lambda_i + \mu_i), P_{i,i-1} = \mu_i/(\lambda_i + \mu_i), i > 0$
- **Linear growth model with immigration** (Ex 6.4, Page 375)
    - $\mu_n = n\mu, n \geq 1; \lambda_n = n\lambda + \theta, n \geq 0.$
    - Deriving M(t)=E[X(t)]
        - Considering M(t+h)=E[X(t+h)]=E[E[X(t+h)|X(t)]]
        - for small h, 3 values for X(t+h) conditional on X(t)
            - X(t)+1, prob=$[\theta + X(t)\lambda]h + o(h)$
            - X(t)-1, prob=$X(t)\mu h + o(h)$
            - X(t), prob = 1- the first two
        - so $E[X(t+h)|X(t)] = X(t) + [\theta + X(t)\lambda - X(t)\mu]h + o(h)$
            - expectation: $M(t+h) = M(t) + (\lambda - \mu)M(t)h + \theta h + o(h)$
            - $(M(t+h) - M(t))/h = (\lambda - \mu)M(t) + \theta + o(h)/h$
            - as $h \to 0, M'(t) = (\lambda - \mu)M(t) + \theta := h(t)$
            - $h'(t) = (\lambda - \mu)M'(t)$ and $h'(t) = (\lambda - \mu)h(t)$
            - $h(t) = Ke^{(\lambda-\mu)t}$ and $\theta + (\lambda - \mu)M(t) = Ke^{(\lambda-\mu)t}$
        - note $M(0) = i$
            - $\lambda \neq \mu$: $M(t) = \theta/(\lambda - \mu)[e^{(\lambda-\mu)t} - 1] + ie^{(\lambda-\mu)t}$
            - $\lambda = \mu$: $M(t) = \theta t + i$
- **A birth-death process with birth rate λ_n and death rate μ_n** with $\mu_0 = 0$
    - T_i denote the *time, starting from state i, it takes for the process to enter state i+1*, $i \geq 0$
    - Deriving $E[T_i]$
        - note $E[T_0] = 1/\lambda_0$ since $\nu_0 = \lambda_0$.
        - Define
            - $I_i = 1$ if the first transition from i is to i+1
            - $I_i = 0$ if the first transition from i is to i-1
            - $\Pr(I_i = 1) = \Pr(\text{birth before death}) = \lambda_i/(\lambda_i + \mu_i)$

        - independent of whether the first transition is from a birth or death, the time until it occurs is exponential with rate $\lambda_i + \mu_i$
        - $E[T_i|I_i = 1] = 1/(\lambda_i + \mu_i)$
            - if the first transition is a birth, then the population size is at i+1, so no additional time is needed
            - => $Var(T_i|I_i = 1) = 1/(\lambda_i + \mu_i)^2$
        - $E[T_i|I_i = 0] = 1/(\lambda_i + \mu_i) + E[T_{i-1}] + E[T_i]$
            - if the first event is death, then the population size becomes i-1,
            - the additional time needed to reach i+1 is equal to the time it takes to return to state i (mean $E[T_{i-1}]$)
            - plus the additional time it then takes to reach i+1 (mean $E[T_i]$)
            - => $Var(T_i|I_i = 0) = 1/(\lambda_i + \mu_i)^2 + Var(T_{i-1}) + Var(T_i)$
        - ==> $E[T_i] = 1/(\lambda_i + \mu_i) + \mu_i/(\lambda_i + \mu_i)(E[T_{i-1}] + E[T_i])$
            - $E[T_i] = 1/\lambda_i + \mu_i/\lambda_i E[T_{i-1}]$
    - expected time to go from state i to state j where i<j
        - $E[T_i] + E[T_{i+1}] + \cdots + E[T_{j-1}]$
    - For $\lambda_n = \lambda, \mu_n = \mu$
        - $E[T_0] = 1/\lambda$
        - $E[T_i] = 1/\lambda + \mu/\lambda E[T_{i-1}] = (1 + \mu E[T_{i-1}])/\lambda$
        - $E[T_1] = (1 + \mu/\lambda)/\lambda, E[T_2] = (1 + \mu/\lambda + [\mu/\lambda]^2)/\lambda$
            - $\lambda \neq \mu$: $E[T_i] = (1 - (\mu/\lambda)^{i+1})/(\lambda - \mu), i \geq 0.$
            - $\lambda = \mu$: $E[T_i] = (i + 1)/\lambda$
        - E[time to go from k to j]=$\sum_{i=k}^{j-1} E[T_i]$, k<j.
            - $\lambda \neq \mu$: $(j - k)/(\lambda - \mu) - (\mu/\lambda)^{k+1}/(\lambda - \mu)[1 - (\mu/\lambda)^{j-k}]/(1 - \mu/\lambda)$
            - $\lambda = \mu$: $(j(j + 1) - k(k + 1)/2/\lambda$
    - Deriving $Var[T_i]$
        - Note $E[T_i|I_i] = 1/(\lambda_i + \mu_i) + (1 - I_i)(E[T_{i-1}] + E[T_i])$
            - $I_i \sim$ Bernoulli$(\lambda_i/(\lambda_i + \mu_i))$
            - => $Var(E[T_i|I_i]) = (E[T_{i-1}] + E[T_i])^2 Var(I_i) = (E[T_{i-1}] + E[T_i])^2 \mu_i \lambda_i/(\mu_i + \lambda_i)^2$
        - $Var(T_i|I_i) = 1/(\lambda_i + \mu_i)^2 + (1 - I_i)[Var(T_{i-1}) + Var(T_i)]$

            - => $E[Var(T_i|I_i)] = 1/(\lambda_i + \mu_i)^2 + \mu_i\lambda_i/(\mu_i + \lambda_i)^2[Var(T_{i-1}) + Var(T_i)]$
        - ==> $Var(T_i) = 1/(\lambda_i(\lambda_i + \mu_i)) + \frac{\mu_i}{\lambda_i} Var(T_{i-1}) + \frac{\mu_i}{\mu_i+\lambda_i} (E[T_{i-1}] + E[T_i])^2$
        - note $Var(T_0) = 1/\lambda_0^2$
        - Var[time to go from k to j]=$\sum_{i=k}^{j-1} Var[T_i]$, k<j.

6.4 The Transition Probability Function

- *transition probabilities* of a continuous-time MC
    - $P_{ij}(t) = \Pr[X(t+s) = j|X(s) = i]$
- **Instantaneous transition rates**: $q_{ij} = \nu_i P_{ij}$
    - q_{ij} is the rate, when in state i, at which the process makes a transition into state j
    - $\nu_i = \sum_j q_{ij}, P_{ij} = q_{ij}/\sum_j q_{ij}$
    - **Lemma 6.2** $\lim_{h\to 0} \frac{1-P_{ii}(h)}{h} = \nu_i, \lim_{h\to 0} \frac{P_{ij}(h)}{h} = q_{ij}, i \neq j$
        - ν_i is the hazard rate of exp dist
        - $1 - P_{ii}(h)$: the probability that a process in state i at time 0 will not be in state i at time h, equals the probability that a transition occurs within time h plus something small compared to h
            - $1 - P_{ii}(h) = \nu_i h + o(h)$
        - $P_{ij}(h)$: the probability that the process goes from state i to state j in a time h, equals the probability that a transition occurs in this time multiplied by the probability that the transition is into state j plus something small compared to h
            - $P_{ij}(h) = h\nu_i P_{ij} + o(h)$
- **Chapman–Kolmogorov equations**: $P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t)P_{kj}(s), \forall s \geq 0, t \geq 0$
    - Condition on $X(t) = k$
    - **Kolmogorov's Backward Equations**: $P'_{ij}(t) = \sum_{k\neq i} q_{ik}P_{kj}(t) - \nu_i P_{ij}(t)$
        - $P_{ij}(h+t) - P_{ij}(t) = \sum_{k\neq i} P_{ik}(h)P_{kj}(t) - [1 - P_{ii}(h)]P_{ij}(t)$
        - taking limit $h \to 0$ and interchange limit and summation
    - **Kolmogorov's Forward Equations**: $P'_{ij}(t) = \sum_{k\neq j} q_{kj}P_{ik}(t) - \nu_j P_{ij}(t)$

- $P_{ij}(t+h) - P_{ij}(t) = \sum_{k \neq j} P_{ik}(t) P_{kj}(h) - [1 - P_{jj}(h)] P_{ij}(t)$
- taking limit $h \to 0$ and interchange limit and summation (under suitable regularity conditions)

- Pure birth process
  - Suppose that the process is presently in state i, and let j > i
    - X_k denote the time the process spends in state k before making a transition into state k+1
    - $\sum_{k=i}^{j-1} X_k$ is the time it takes until the process enters state j
    - $X(t) < j \leftrightarrow \sum_{k=i}^{j-1} X_k > t$
    - $\Pr[X(t) < j | X(0) = i] = \Pr[\sum_{k=i}^{j-1} X_k > t]$
      - X_k are independent exps with rate λ_k
      - sum of independent exps with distinct rates (*hypoexponential dist*)
      - = $\sum_{k=i}^{j-1} e^{-\lambda_k t} \prod_{r \neq k, r=i}^{j-1} \frac{\lambda_r}{\lambda_r - \lambda_k}$ (assuming $\lambda_i \neq \lambda_j, \forall i \neq j$)
    - $\Pr[X(t) = j | X(0) = i] = \Pr[X(t) < j+1 | X(0) = i] - \Pr[X(t) < j | X(0) = i]$
    - $P_{ii}(t) = \Pr(X_i > t) = e^{-\lambda_i t}$
  - Backward equation
    - $P_{ij}{}'(t) = \lambda_i [P_{i+1,j}(t) - P_{ij}(t)]$
  - Forward equation
    - $P_{ii}{}'(t) = -\lambda_i P_{ii}(t)$
      - $P_{ii}(t) = e^{-\lambda_i t}$
    - $P_{ij}{}'(t) = \lambda_{j-1} P_{i,j-1}(t) - \lambda_j P_{ij}(t)$
      - $P_{ij}(t) = \lambda_{j-1} e^{-\lambda_j t} \int_0^t e^{\lambda_j s} P_{i,j-1}(s) ds$
      - treating $e^{\lambda_j t} P_{ij}(t)$ as one function
- Birth and death process
  - Backward equation
    - $P_{0j}{}'(t) = \lambda_0 [P_{1j}(t) - P_{0j}(t)]$
    - $P_{ij}{}'(t) = \lambda_i P_{i+1,j}(t) + \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t)$
  - Forward equation
    - $P_{i0}{}'(t) = \mu_1 P_{i1}(t) - \lambda_0 P_{i0}(t)$

- $P_{ij}'(t) = \lambda_{j-1} P_{i,j-1}(t) + \mu_{j+1} P_{i,j+1}(t) - (\lambda_j + \mu_j) P_{ij}(t)$

- **Continuous-Time Markov Chain Consisting of Two States (Example 6.11, Page 386)**
  - A machine works for an exponential amount of time having mean $1/\lambda$ before breaking down; and it takes an exponential amount of time having mean $1/\mu$ to repair the machine
    - A birth-death process: $\lambda_0 = \lambda, \mu_1 = \mu, \lambda_i = 0, i \neq 0, \mu_i = 0, i \neq 1$
    - $P_{00}(0) = 1, P_{10}(0) = 0$
    - $P_{00}(t) = \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu}$
    - $P_{10}(t) = -\frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu}$
    - some results
      - continuous-time MC: $\nu_0 = \lambda, \nu_1 = \mu$. $\lim_{t\to 0} P_{10}(t)/t = \mu, \lim_{t\to 0} P_{01}(t)/t = \lambda$
      - embeded discrete MC: transition prob $P_{10} = P_{01} = 1$; stationary dist $\pi_0 = \pi_1 = 0.5$
      - limiting prob of continuous-time MC: $\lim_{t\to\infty} P_{10}(t) = \mu/(\mu + \lambda)$

6.5 Limiting Probabilities

- $P_{ij}(t)$ often converges to a limiting value that is independent of the initial state i.
  - let $P_j = \lim_{t\to\infty} P_{ij}(t), \sum_j P_j = 1$ (*stationary probabilities*)
    - $P_{ij}(t)$ will be constant if we start from the limiting prob.
  - Forward equation: $P'_{ij}(t) = \sum_{k\neq j} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t)$
    - $\lim_{t\to\infty} P'_{ij}(t) = \sum_{k\neq j} q_{kj} P_k - \nu_j P_j$
      - note $\lim_{t\to\infty} P'_{ij}(t) = 0$ (*if not, $P_{ij}(t)$ will go to infinity*)
  - So $\nu_j P_j = \sum_{k\neq j} q_{kj} P_k$ (*balance equations*)
    - Sufficient condition for P_j existing (*ergodic chain*)
      - all states of the MC communicate in the sense that starting in state i there is a positive probability of ever being in state j, for all i, j and
      - the MC is positive recurrent in the sense that, starting in any state, the mean time to return to that state is

finite

- P_j being the long-run proportion of time that the process is in state j
- $\nu_j P_j$ = rate at which the process leaves state j
  - when the process is in state j, it leaves at rate ν_j
  - P_j is the proportion of time it is in state j
- $\sum_{k \neq j} q_{kj} P_k$ = rate at which the process enters state j
  - $q_{kj} P_j$ = rate at which transitions from k to j occur
- hence equality of the rates at which the process enters and leaves state j

- Birth and death process
  - birth rate λ_n, death rate μ_n
    - state 0: leaving rate $\lambda_0 P_0$; entering rate $\mu_1 P_1$
    - state $n \geq 1$: leaving rate $(\lambda_n + \mu_n) P_n$; entering rate $\mu_{n+1} P_{n+1} + \lambda_{n-1} P_{n-1}$
    - $\Rightarrow \lambda_n P_n = \mu_{n+1} P_{n+1}, n \geq 0$
    - hence $P_n = \frac{\prod_{i=1}^{n} \lambda_{i-1}}{\prod_{i=1}^{n} \mu_i} P_0$, and $P_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \frac{\prod_{i=1}^{n} \lambda_{i-1}}{\prod_{i=1}^{n} \mu_i}}$
      - limiting prob exists iff $\sum_{n=1}^{\infty} \frac{\prod_{i=1}^{n} \lambda_{i-1}}{\prod_{i=1}^{n} \mu_i} < \infty$
  - linear growth model with immigration
    - $\lambda_n = n\lambda + \theta, \mu_n = n\mu$
    - previous condition reduces to $\sum_{n=1}^{\infty} \frac{\prod_{i=1}^{n} (\theta + (i-1)\lambda)}{n! \mu^n} < \infty$
      - consider the limit of ratio of two adjacent terms: $\lim_n \frac{\theta + n\lambda}{(n+1)\mu} = \frac{\lambda}{\mu}$
      - therefore when $\lambda < \mu$, the limiting prob exists!
  - a birth-death process with $\mu_n = \mu, n \geq 1, \lambda_n = (M-n)\lambda, n \leq M, \lambda_n = 0, n > M$
    - *Continuous-Time Markov Chain Consisting of Finite States (Ex 6.13, Page 393)*
    - $P_n = \frac{(\lambda/\mu)^n M!/(M-n)!}{1 + \sum_{n=1}^{M} (\lambda/\mu)^n M!/(M-n)!}, n = 0, \ldots, M$

6.6 Time Reversibility

- For a continuous-time ergodic MC with limiting P_j
  - consider its discrete-time transition matrix P_{ij}
    - its limiting prob π_i: $\pi_i = \sum_j \pi_j P_{ji}$
  - Claim: $P_i = \pi_i/\nu_i/(\sum_j \pi_j/\nu_j)$
    - $\nu_i P_i = \sum_{j \neq i} P_j q_{ji}$
    - note $P_{ii} = 0$
  - Going backward in time, the amount of time the process spends in state i is also exp with the same rate
    - assume the chain is already in limiting dist
    - Pr(process is in state i throughout $[t-s,t]|X(t)=i)$
      - = Pr(process is in state i throughout $[t-s,t]) / \Pr(X(t)=i) = \Pr(X(t-s)=i)e^{-\nu_i s} / \Pr(X(t)=i) = e^{-\nu_i s}$
    - sequence of states visited by the reversed process constitutes a discrete-time MC with transition probabilities $Q_{ij} = \pi_j P_{ji}/\pi_i$
    - reversed process is a continuous-time MC with the same transition rates as the forward-time process and with one-stage transition probabilities Q_{ij}
  - Continuous-time MC is *time reversible* if $\pi_i P_{ij} = \pi_j P_{ji}$
    - time reversible, in the sense that the process reversed in time has the same probabilistic structure as the original process, when the embedded chain is time reversible
    - <=> $P_i q_{ij} = P_j q_{ji}$
    - P_i is the proportion of time in state i and q_{ij} is the rate when in state i that the process goes to j
    - the rate at which the process goes directly from state i to state j is equal to the rate at which it goes directly from j to i
- **Proposition 6.5** An ergodic birth and death process is time reversible
  - the rate at which a birth and death process goes from state i to state i+1 is equal to the rate at which it goes from i+1 to i
  - In any length of time t the number of transitions from i to i+1 must equal to within 1 the number from i+1 to i
    - since between each transition from i to i+1 the process must return to i, and this can only occur through i+1, and

vice versa.

- **Proposition 6.7** If for some $\{P_i \geq 0\}$, $\sum_i P_i = 1$, and $P_i q_{ij} = P_j q_{ji}, \forall i \neq j$, then the continuous-time MC is time reversible and P_i is the limiting probability of being in state i.
    - $\sum_{j \neq i} P_i q_{ij} = \sum_{j \neq i} P_j q_{ji}$
    - note $\nu_i = \sum_{j \neq i} q_{ij}$
    - so $v_i P_i = \sum_{j \neq i} P_j q_{ji}$ (*balance equation*)
- **Proposition 6.8** A time reversible chain with limiting prob $P_j, j \in S$ that is truncated to the set $A \in S$ and remains irreducible is also time reversible and has limiting probabilities P_j^A given by $P_j^A = \frac{P_j}{\sum_{i \in A} P_i}, j \in A$.

6.7 Uniformization

- Consider a continuous-time Markov chain with $\nu_i = \nu, \forall i$
    - N(t) denote the number of state transitions by time t, then $\{N(t), t \geq 0\}$ will be a Poisson process with rate ν.
        - N(t) ~ Poisson(νt)
    - $P_{ij}(t) = \Pr(X(t) = j | X(0) = i) = \sum_{n=0}^{\infty} \Pr(N(t) = n | X(0) = i) \Pr(X(t) = j | X(0) = i, N(t) = n)$
        - $\Pr(N(t) = n | X(0) = i) = e^{-\nu t}(\nu t)^n / n!$ (Poisson rv)
        - $\Pr(X(t) = j | X(0) = i, N(t) = n) = P_{ij}^n$
            - knowing that N(t) = n gives us no information about which states were visited, since the distribution of time spent in each state is the same for all states
            - P_{ij}^n is the n-stage transition probability associated with the discrete-time MC with transition prob P_{ij}
    - So $P_{ij}(t) = \sum_{n=0}^{\infty} P_{ij}^n e^{-\nu t}(\nu t)^n / n!$
        - approximate $P_{ij}(t)$ by taking a partial sum and then computing the n stage probabilities P_{ij}^n.
- uniformizing the rate in which a transition occurs from each state by introducing transitions from a state to itself
    - assume $\nu_i \leq \nu, \forall i$
    - Define P_{ij}^*: $1 - \nu_i / \nu$ for $j = i$ and $\nu_i / \nu P_{ij}$ for $j \neq i$.
    - $P_{ij}(t) = \sum_{n=0}^{\infty} P_{ij}^{*n} e^{-\nu t}(\nu t)^n / n!$

6.8 Computing the Transition Probabilities

- Define r_{ij}: q_{ij} if $i \neq j$, and $-\nu_i$ if $i = j$.
  - Kolmogorov backward equations: $P'_{ij}(t) = \sum_k r_{ik} P_{kj}(t)$
  - Kolmogorov forward equations: $P'_{ij}(t) = \sum_k r_{kj} P_{ik}(t)$
- Define corresponding matrix $\mathbf{R}, \mathbf{P}(t), \mathbf{P}'(t)$
  - $\mathbf{P}'(t) = \mathbf{R}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{R}$
    - so $\mathbf{P}(t) = \mathbf{P}(0)e^{\mathbf{R}t} = e^{\mathbf{R}t}$, since $\mathbf{P}(0) = \mathbf{I}$.
    - $e^{\mathbf{R}t} = \sum_{n=0}^{\infty} \mathbf{R}^n t^n / n!$
- Approximation of $\mathbf{P}(t)$
  - $e^{\mathbf{R}t} = \lim_{n\to\infty}(\mathbf{I} + \mathbf{R}t/n)^n$
    - taking $n = 2^k$: approximating by matrix square (multiplication)
    - for large n, $\mathbf{I} + \mathbf{R}t/n$ all nonnegative elements
  - $e^{-\mathbf{R}t} = \lim_{n\to\infty}(\mathbf{I} - \mathbf{R}t/n)^n$
    - $(\mathbf{I} - \mathbf{R}t/n)^{-1}$ nonnegative elements