Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE550 Stochastic Processes and Bayesian Inference

Exam 2019, January 16, 8:30 - 12:30
Allowed aids: Chalmers-approved calculator.
Total number of points: 30. To pass, at least 12 points are needed
There is an appendix containing relevant information about some probability distributions.

NOTE: By request, there is an appendix containing a Swedish version of all questions. You may choose to answer in English or Swedish.

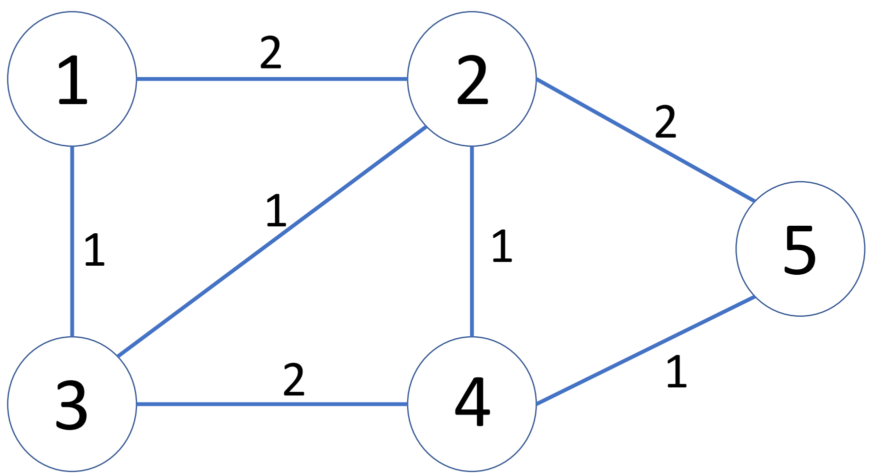

Figure 1: The graph for question 1.

1. (5 points) Consider the random walk on the weighted undirected graph given in Figure 1.
   (a) Write down the transition matrix for the corresponding discrete-time Markov chain.
   (b) Compute the stationary distribution for the random walk.
   (c) Define time reversibility for a discrete irreducible Markov chain with transition matrix P. Prove that any random walk on a weighted undirected graph is time reversible.
2. (4 points) Assume $x \mid p \sim \text{Geometric}(p)$, so that x has a Geometric distribution with parameter p.

(a) Assume that the prior is $p \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ for some parameters $\alpha > 0$ and $\beta > 0$. Compute the posterior distribution $p \mid x$ and find its name and parameters.

(b) Find a function f so that we have the following for the marginal probability mass function $\pi(x)$:

$$\pi(x) = Cf(x)$$

for some constant C not depending on x.

3. (4 points) Let a discrete-time Markov chain on the state space $\{1, 2\}$ have transition matrix

$$P = \begin{bmatrix} 1-p & p \\ 0 & 1 \end{bmatrix}$$

for some p with $0 < p < 1$.

(a) Compute the fundamental matrix.

(b) For a chain starting in 1, what is the expected number of steps until absorbtion?

(c) For a chain starting in 1, what is the variance of the number of steps until absorbtion?

4. (5 points) Assume a branching process is such that the probability of zero, one, or two offspring is 1/8, 1/2, or 3/8, respectively.

(a) What is the expected size of Z_5, the fifth generation of the branching process?

(b) Write down the probability generating function for the offspring process[1].

(c) Compute the probability that the process will eventually go extinct.

5. (2 points) Consider the probability density on the real numbers defined by

$$\pi(x) = C \exp\left(-(x + \sin x)^2\right)$$

where C is a constant. Describe in detail how to use the Metropolis Hastings algorithm to obtain an approximate sample from a random variable with this density. Use as proposal function $q(x \mid x_0) = \text{Normal}(x; x_0, \sigma_a^2)$, where x_0 is the previous value in the chain and σ_a^2 is some constant.

6. (4 points) Assume $\{N_t\}_{t \geq 0}$ is a Poisson process with parameter λ, and let S_n be the arrival time of the n'th observation.

(a) Find the expectation and variance of $S_m - S_n$, when $m > n$.

(b) Compute the correlation $\text{corr}(S_m, S_n)$, when $m > n$.

[1] In the original exam, the question was formulated "Write down the probability generating function".

7. (4 points) Customers arrive to a drop-in hair salon as a Poisson process with the rate of 3 per hour. Two people work at the salon: The owner A and her assistant B. When a customer arrives, the following happens: If either the assistant or the owner is working, the other one takes the new customer, but if none of them are working, the assistant is set to work. If they are both working, the customer will wait in the single waiting chair. If the waiting chair is filled, the new customer goes away.

   The length of a haircut done by the owner is exponentially distributed with expectation 20 minutes, while the length of a haircut done by the assistant is exponentially distributed with expectation 30 minutes. When they finish and there is a customer waiting, the person who finishes will immediately start working on this customer.

   (a) Write down the generator matrix Q for the 5 possible "states" of the hair salon.

   (b) We would like to compute the long term[2] proportion of time that the owner is working but not the assistant. Write down exactly which equations to solve, and how to make this computation without a computer; however you do not need to do the actual computation.

8. (2 points) Assume A is a real 3×3 matrix with the property that there exists an invertible matrix S such that $A = SDS^{-1}$ with D a diagonal matrix with values $1, 1/2, 1/3$ along the diagonal. Compute the determinant of the matrix e^A.

[2] In the original exam, the formulation was "expected" instead of "long term"

Appendix: Some probability distributions

The Bernoulli distribution

If $x \in \{0, 1\}$ has a Bernoulli distribution with parameter $0 \le p \le 1$, then the probability mass function is

$$\pi(x) = p^x(1-p)^{1-x}.$$

We write $x \mid p \sim \text{Bernoulli}(p)$ and $\pi(x \mid p) = \text{Bernoulli}(x; p)$.

The Beta distribution

If $x \in [0, 1]$ has a Beta distribution with parameters with $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}.$$

We write $x \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Beta}(x; \alpha, \beta)$.

The Beta-Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Beta-Binomial distribution, with n a positive integer and parameters $\alpha > 0$ and $\beta > 0$, then the probability mass function is

$$\pi(x \mid n, \alpha, \beta) = \binom{n}{x} \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n+\alpha+\beta)}.$$

We write $x \mid n, \alpha, \beta \sim \text{Beta-Binomial}(n, \alpha, \beta)$ and $\pi(x \mid n, \alpha, \beta) = \text{Beta-Binomial}(x; n, \alpha, \beta)$.

The Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Binomial distribution, with n a positive integer and $0 \le p \le 1$, then the probability mass function is

$$\pi(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}.$$

We write $x \mid n, p \sim \text{Binomial}(n, p)$ and $\pi(x \mid n, p) = \text{Binomial}(x; n, p)$.

The Dirichlet distribution

If $x = (x_1, x_2, \ldots, x_n)$ has a Dirichlet distribution, with $x_i \ge 0$ and $\sum_{i=1}^{n} x_i = 1$ and with parameters $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_1 > 0, \ldots, \alpha_n > 0$, then the density function is

$$\pi(x \mid \alpha) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)} p_1^{\alpha_1-1} p_2^{\alpha_2-1} \cdots p_n^{\alpha_n-1}.$$

We write $x \mid \alpha \sim \text{Dirichlet}(\alpha)$ and $\pi(x \mid \alpha) = \text{Dirichlet}(x; \alpha)$.

The Exponential distribution

If $x \geq 0$ has an Exponential distribution with parameter $\lambda > 0$, then the density is

$$\pi(x \mid \lambda) = \lambda \exp(-\lambda x)$$

We write $x \mid \lambda \sim \text{Exponential}(\lambda)$ and $\pi(x \mid \lambda) = \text{Exponential}(x; \lambda)$. The expectation is $1/\lambda$ and the variance is $1/\lambda^2$.

The Gamma distribution

If $x > 0$ has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

We write $x \mid \alpha, \beta \sim \text{Gamma}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Gamma}(x; \alpha, \beta)$.

The Geometric distribution

If $x \in \{1, 2, 3, \ldots\}$ has a Geometric distribution with parameter $p \in (0, 1)$, the probability mass function is

$$\pi(x \mid p) = p(1-p)^{x-1}$$

We write $x \mid p \sim \text{Geometric}(p)$ and $\pi(x \mid p) = \text{Geometric}(x; p)$. The expectation is $1/p$ and the variance $(1-p)/p^2$.

The Normal distribution

If the real x has a Normal distribution with parameters μ and σ^2, its density is given by

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

We write $x \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$ and $\pi(x \mid \mu, \sigma^2) = \text{Normal}(x; \mu, \sigma^2)$.

The Poisson distribution

If $x \in \{0, 1, 2, \ldots\}$ has Poisson distribution with parameter $\lambda > 0$ then the probability mass function is

$$e^{-\lambda} \frac{\lambda^x}{x!}.$$

We write $x \mid \lambda \sim \text{Poisson}(\lambda)$ and $\pi(x \mid \lambda) = \text{Poisson}(x; \lambda)$.

Appendix: Frågorna på svenska

1. (5 poäng) Observera slumpvandringen som defineras av Figur 1.
   (a) Skriv ner övergångsmatrisen för motsvarande tids-diskreta Markovkedja.
   (b) Beräkna stationärfördelningen för slumpvandringen.
   (c) Definera tids-reversibilitet för en diskret irreducibel Markovkedja med övergångsmatris P. Bevisa att en slumpvandring på en viktad oriktad graf är tids-reversibel.

2. (4 poäng) Anta $X \mid p \sim \text{Geometric}(p)$, så att X har en geometrisk föredelning med parameter p.
   (a) Anta apriori-fördelningen $p \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ för parmetrar $\alpha > 0$ och $\beta > 0$. Beräkna posteriori-fördelningen $p \mid X$ och hitta dens namn och parametrar.
   (b) Hitta en funktion f så att följande gäller för den marginella sannolikhetsfunktionen $\pi(x)$:
   $$\pi(x) = Cf(x)$$
   för någon konstant C som inte beror på x.

3. (4 poäng) Låt en tids-diskret Markovkedja med tillstånsrum $\{1, 2\}$ ha övergångsmatris
   $$P = \begin{bmatrix} 1-p & p \\ 0 & 1 \end{bmatrix}$$
   för någon p med $0 < p < 1$.
   (a) Beräkna den fundamentala matrisen.
   (b) För en kedja som startar i 1, vad är förväntad antal steg till absorption?
   (c) För en kedja som startar i 1, vad är variansen av antal steg till absorption?

4. (5 poäng) Anta en förgreningsprocess är sådan att sannolikheten för noll, ett eller två avkom är 1/8, 1/2 respektiva 3/8.
   (a) Vad är förväntad storlek av Z_5, femte generation av förgreningsprocessen?
   (b) Skriv ner den sannolikhetsgenererande funktionen för fördelningen för avkommor[3].
   (c) Beräkna sannolikheten att processen dör ut.

5. (2 poäng) Observera en stokastisk variabel med reella värden och täthetsfunktion
   $$\pi(x) = C \exp\left(-(x + \sin x)^2\right)$$
   där C är en konstant. Beskriva detaljerad hur man kan använda Metropolis-Hastings algoritmen för att ta fram ett approximativt stickprov från en slumpvariabel med denna täthetsfunktion. Använd förslagsfunktionen $q(x \mid x_0) = \text{Normal}(x; x_0, \sigma_a^2)$, där x_0 är kedjans tidigare värde och σ_a^2 är någon konstant.

[3] I originaltentan så formulerades frågan som “Skriv ner den sannolikhetsgenererande funktionen.

6. (4 poäng) Anta $\{N_t\}_{t\geq 0}$ är en Poissonprocess med parameter λ och låt S_n vara ankomsttiden för observation nummer n.

   (a) Beräkna väntevärde och varians för $S_m - S_n$ när $m > n$.

   (b) Beräkna korrelationen $\mathrm{corr}(S_m, S_n)$ när $m > n$.

7. (4 poäng) Kunder ankommer en drop-in frisörsalong enligt en Poissonprocess med rate 3 per timme. Två personer arbetar i salongen: Ägaren A och hennes assistent B. När en ny kund ankommer så händer följande: Om antingen assistenten eller ägaren arbetar så börjar den som inte arbetar att arbeta med kunden. Om vare sig ägaren eller assistenten arbetar så börjar assistenten arbeta med kunden. Om både ägaren och assistenten arbetar så väntar kunden i salongens ända stol för väntande kunder. Om det redan finns en kund i denna stol så lämnar den nya kunden salongen.

   Tiden det tar ägaren att klippa är exponensialfördelad med väntevärde 20 minuter, medan tiden det tar assistenten att klippa är exponensialfördelad med väntevärde 30 minuter. Om en klippning avslutas medan det finns någon som väntar så börjar den som har klippat genast att arbeta med den väntande kunden.

   (a) Skriv ner generatormatrisen Q för de 5 möjliga tillstånd som frisörsalongen kan vara i.

   (b) Vi vill gärna beräkna andelen tid, i det långa loppet[4], under vilken ägaren arbetar medan assitenten inte arbetar. Skriv ner exakt vilka ekvationer som behöver lösas, och hur man gör beräkningarna utan dator. Dock behöver du inte genomföra beräkningarna.

8. (2 poäng) Anta A är en reell 3×3 matris så att det existerar en invertibel matris S så att $A = SDS^{-1}$ där D är en diagonalmatris med värden $1, 1/2, 1/3$ långsmed diagonalen. Beräkna determinanten av matrisen e^A.

[4] I den ursprungliga tentan var formuleringen “väntevärdet för andelen tid”

Petter Mostad
Applied Mathematics and Statistics
Chalmers

Suggested solutions for
MVE550 Stochastic Processes and Bayesian Inference
Exam 2019, January 16

1. (a) We get

$$P = \begin{bmatrix} 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

(b) For any random walk on a weighted undirected graph, the stationary distribution can be read off the graph. Note that the sum of all weights is 10. Thus

$$\begin{aligned} v &= \left(\frac{1+2}{2\cdot 10}, \frac{2+1+1+2}{2\cdot 10}, \frac{1+1+2}{2\cdot 10}, \frac{1+1+2}{2\cdot 10}, \frac{1+2}{2\cdot 10}\right) \\ &= (0.15, 0.3, 0.2, 0.2, 0.15) \end{aligned}$$

(c) Let v be the probability vector representing the stationary distribution. Then the chain is time-reversible if and only if, for all i and j,

$$v_i P_{ij} = v_j P_{ji}.$$

Let w_{ij} denote the weight on the line connecting state i and state j. Then, for the random walk, we have

$$P_{ij} = \frac{w_{ij}}{\sum_k w_{ik}}$$

and

$$v_i = \frac{\sum_k w_{ik}}{\sum_s \sum_k w_{sk}}.$$

Thus

$$\begin{aligned} v_i P_{ij} &= \frac{\sum_k w_{ik}}{\sum_s \sum_k w_{sk}} \cdot \frac{w_{ij}}{\sum_k w_{ik}} = \frac{w_{ij}}{\sum_s \sum_k w_{sk}} \\ v_j P_{ji} &= \frac{\sum_k w_{jk}}{\sum_s \sum_k w_{sk}} \cdot \frac{w_{ji}}{\sum_k w_{jk}} = \frac{w_{ji}}{\sum_s \sum_k w_{sk}}. \end{aligned}$$

As $w_{ij} = w_{ji}$ for all i and j, we have time-reversibility.

2. (a) We get

$$\begin{aligned}\pi(p \mid x) &\propto_p \pi(x \mid p)\pi(p) \\ &\propto_p \text{Geometric}(x; p) \cdot \text{Beta}(p; \alpha, \beta) \\ &\propto_p p(1-p)^{x-1} p^{\alpha-1}(1-p)^{\beta-1} \\ &\propto_p p^{\alpha+1-1}(1-p)^{\beta+x-1-1}\end{aligned}$$

Thus $p \mid x \sim \text{Beta}(\alpha + 1, \beta + x - 1)$.

(b) We get

$$\begin{aligned}\pi(x) &= \frac{\pi(x \mid p)\pi(p)}{\pi(p \mid x)} \\ &\propto_x \frac{\text{Geometric}(x; p)}{\text{Beta}(p; \alpha + 1, \beta + x - 1)} \\ &\propto_x \frac{p(1-p)^{x-1}}{\frac{\Gamma(\alpha+1+\beta+x-1)}{\Gamma(\alpha+1)\Gamma(\beta+x-1)} p^{\alpha+1-1}(1-p)^{\beta+x-1-1}} \\ &\propto_x \frac{\Gamma(\beta + x - 1)}{\Gamma(\beta + x + \alpha)}\end{aligned}$$

Thus $f(x) = \frac{\Gamma(\beta+x-1)}{\Gamma(\beta+x+\alpha)}$. When α is an integer, this corresponds to $f(x) = \frac{1}{(\beta+x-1)(\beta+x)\cdots(\beta+x+\alpha-1)}$.

3. (a) The fundamental matrix is

$$F = (I - Q)^{-1} = (1 - (1 - p))^{-1} = p^{-1} = 1/p$$

i.e., the 1×1 matrix with the single element $1/p$.

(b) The expected number of steps until absorbtion can be found from the fundamental matrix, i.e., it is $1/p$.

(c) Let X denote the number of steps until absorbtion. Then we can read from the definition of P that, for $k = 1, 2, \ldots,$

$$\mathrm{P}(X = k) = p(1 - p)^{k-1}.$$

This means that $X \sim \text{Geometric}(p)$. From the appendix we have that $\text{Var}[X] = (1 - p)/p^2$, so this is the answer.

4. (a) The offspring distribution has expectation

$$0 \cdot \frac{1}{8} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{8} = \frac{5}{4}$$

Thus

$$\mathrm{E}(Z_5) = \left(\frac{5}{4}\right)^5 = \frac{3125}{1024} = 3.051758$$

(b) We get

$$G(s) = \mathrm{E}(s^X) = \frac{1}{8} + \frac{1}{2}s + \frac{3}{8}s^2$$

(c) We get

$$\begin{aligned} G(s) &= s \\ \frac{1}{8} + \frac{1}{2}s + \frac{3}{8}s^2 &= s \\ 3s^2 - 4s + 1 &= 0 \\ (s-1)(3s-1) &= 0 \end{aligned}$$

Thus the smallest positive root of $G(s) = s$ is 1/3, which is the probability of extinction.

5. The algorithm starts with selecting an initial real value $x^{(0)}$. Then, for $i = 1, 2, \ldots,$, the algorithm generates $x^{(i)}$ as follows:

- Simulate a proposed value $y \sim \text{Normal}(x^{(i-1)}, \sigma_a^2)$.
- Compute the acceptance probability:

$$\begin{aligned} p &= \min\left(1, \frac{\pi(y)q(x^{(i-1)} \mid y)}{\pi(x^{(i-1)})q(y \mid x^{(i-1)})}\right) \\ &= \min\left(1, \exp\left(-(y + \sin y)^2 + (x^{(i-1)} + \sin x^{(i-1)})^2\right)\right). \end{aligned}$$

  NOTE: The quotient above is *not* $\frac{\pi(y)q(y|x^{(i-1)})}{\pi(x^{(i-1)})q(x^{(i-1)}|y)}$.

- With probability p, set $x^{(i)} = y$, otherwise, set $x^{(i)} = x^{(i-1)}$.

The distribution of the sequence $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$ will now converge to a distribution with density $\pi(x)$.

6. For $i = 1, 2, \ldots,$ let X_i be the holding time between arrival $i-1$ and arrival i. Then all the X_i are independent and $X_i \sim \text{Exponential}(\lambda)$. Also $S_n = \sum_{i=1}^n X_i$ and $S_m - S_n = \sum_{i=n+1}^m X_i$.

(a) We get

$$\begin{aligned} \mathrm{E}(S_m - S_n) &= \mathrm{E}\left(\sum_{i=n+1}^m X_i\right) = \sum_{i=n+1}^m \mathrm{E}(X_i) = \frac{m-n}{\lambda} \\ \mathrm{var}(S_m - S_n) &= \mathrm{var}\left(\sum_{i=n+1}^m X_i\right) = \sum_{i=n+1}^m \mathrm{var}(X_i) = \frac{m-n}{\lambda^2}. \end{aligned}$$

(b) We get

$$\begin{aligned}
\operatorname{corr}(S_m, S_n) &= \frac{\operatorname{cov}(S_m, S_n)}{\sqrt{\operatorname{var}(S_m)\cdot\operatorname{var}(S_n)}} \\
&= \frac{\operatorname{cov}\left(S_n + \sum_{i=n+1}^m X_i, S_n\right)}{\sqrt{\operatorname{var}(S_m)\cdot\operatorname{var}(S_n)}} \\
&= \frac{\operatorname{cov}\left(S_n, S_n\right) + \operatorname{cov}\left(\sum_{i=n+1}^m X_i, S_n\right)}{\sqrt{\operatorname{var}(S_m)\cdot\operatorname{var}(S_n)}} \\
&= \frac{\operatorname{cov}(S_n, S_n)}{\sqrt{\operatorname{var}(S_m)\cdot\operatorname{var}(S_n)}} \\
&= \frac{\sqrt{\operatorname{var}(S_n)}}{\sqrt{\operatorname{var}(S_m)}} \\
&= \sqrt{\frac{n/\lambda^2}{m/\lambda^2}} = \sqrt{\frac{n}{m}}.
\end{aligned}$$

7. (a) Ordering the states of the hair salon as

   i. No customers.
   ii. Only A working.
   iii. Only B working.
   iv. Both A and B working but no-one waiting.
   v. Both A and B working and one person waiting.

   we get (using hours as the unit of time)

$$Q = \begin{bmatrix} -3 & 0 & 3 & 0 & 0 \\ 3 & -6 & 0 & 3 & 0 \\ 2 & 0 & -5 & 3 & 0 \\ 0 & 2 & 3 & -8 & 3 \\ 0 & 0 & 0 & 5 & -5 \end{bmatrix}.$$

(b) Let $v = (v_1, v_2, \ldots, v_5)$ be the stationary distribution for the process. Then the answer to the question is given by v_2. We know that $vQ = 0$ and that $\sum_{i=1}^5 v_i = 1$. These equations represent 6 equations for the 5 unknown components of v. The equations are

$$\begin{aligned}
-3v_1 + 3v_2 + 2v_3 &= 0 \qquad (1)\\
-6v_2 + 2v_4 &= 0 \qquad (2)\\
3v_1 - 5v_3 + 3v_4 &= 0 \qquad (3)\\
3v_2 + 3v_3 - 8v_4 + 5v_5 &= 0 \qquad (4)\\
3v_4 - 5v_5 &= 0 \qquad (5)\\
v_1 + v_2 + v_3 + v_4 + v_5 &= 1 \qquad (6)
\end{aligned}$$

To find v_2 we need to solve this system. We may in fact remove any of the equations (1) through (5).

8. From the definition of the exponential matrix and using $A = SDS^{-1}$, we get

$$e^A = S e^D S^{-1}$$

Thus

$$\begin{aligned} \det(e^A) &= \det(S e^D S^{-1}) = \det(S)\det(e^D)\det(S^{-1}) \\ &= \det(S)\det\begin{bmatrix} e^1 & 0 & 0 \\ 0 & e^{1/2} & 0 \\ 0 & 0 & e^{1/3} \end{bmatrix}\det(S)^{-1} \\ &= e^1 e^{1/2} e^{1/3} = e^{11/6} = 6.254701. \end{aligned}$$