Student Name: Erin Sarlak

Collaboration Statement:

Total hours spent: 8

I discussed ideas with these individuals:

- Patrick Feeney
- ...

I consulted the following resources:

- Data Science Basics Youtube Channel
- Textbook
- ...

By submitting this assignment, I affirm this is my own original work that abides by the course collaboration policy.

Links: [HW3 instructions] [collab. policy]

Contents

1a: Problem Statement

Define $\Sigma = LL^T$. Show the following:

$$|\det(L^{-1})| = \frac{1}{(\det\Sigma)^{\frac{1}{2}}} \tag{1}$$

1a: Solution

$$= \frac{1}{(\det\Sigma)^{\frac{1}{2}}} \tag{2}$$

Substitute $\Sigma = LL^T$ in expression 2.

$$= \frac{1}{(\det(LL^T))^{\frac{1}{2}}} \tag{3}$$
$$= \frac{1}{(\det(L)\cdot\det(L^T))^{\frac{1}{2}}} \tag{4}$$

The determinant of a triangular matrix is the product of its diagonal entries. Because the diagonal in the matrix remains the same, $\det(L) = \det(L^T)$.

$$= \frac{1}{((\det(L))^2)^{\frac{1}{2}}} \tag{5}$$
$$= \frac{1}{\det(L)} \tag{6}$$
$$= |\det(L^{-1})| \tag{7}$$
$$\tag{8}$$

1b: Problem Statement

Show that the pdf of x is given by:

$$p(x) = \frac{1}{(2\pi)^{\frac{D}{2}}}\frac{1}{(\det\Sigma)^{\frac{1}{2}}}e^{-\frac{1}{2}(x-\mu)^T\Sigma^1(x-\mu)} \tag{9}$$

1b: Solution

$$p(x) = f(S(x)) \cdot |\det(J_s(x))| \tag{10}$$

Substitute $|\det(J_s(x))| = |\det(L^{-1})| = \frac{1}{\det(\Sigma)^{\frac{1}{2}}}$ in expression 10.

$$= f(S(x)) \cdot \frac{1}{\det(\Sigma)^{\frac{1}{2}}} \tag{11}$$

Recall that $f(x) = \frac{1}{(2\pi)^{\frac{D}{2}}} \cdot e^{\frac{-1}{2}x^T x}$. Then, we can substitute in f(x), $x = S(x) = L^{-1}(m - x)$.

$$= \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{(\det\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(L^{-1}(m-x))^T (L^{-1}(m-x))} \tag{12}$$

Let's simplify the part in red. Because $(AB)^T = B^T A^T$ (matrix derivatives cheat sheet), $(L^{-1}(m - x))^T (L^{-1}(m - x)) = (x - m)^T (L^{-1}) L^{-1}(x - m)$. Substituting $\Sigma^{-1} = (L^{-1})^T L^{-1}$ results in $(x - m)^T \Sigma^{-1}(x - m)$. Then, expression 12 becomes the following, where m is μ.

$$p(x) = \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{(\det\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(x-m)^T \Sigma^{1}(x-m)} \tag{13}$$

1c: Problem Statement

Complete the Python code below, to show how to turn samples from a standard 1D Gaussian, via NumPy's 'randn()' into a sample from a multivariate Gaussian.x

1c: Solution

```
import numpy as np

def sample_from_mv_gaussian(mu_D, Sigma_DD, random_state=np.random):
  ''' Draw sample from multivariate Gaussian

  Args
  ----
  mu_D : 1D array, size D
    Mean vector
  Sigma_DD : 2D array, shape (D, D)
    Covariance matrix. Must be symmetric and positive definite.

  Returns
  -------
  x_D : 1D array, size D
    Sampled value of Gaussian with provided mean and covariance
  '''
  D = mu_D.size
  L_DD = np.linalg.cholesky(Sigma_DD) # compute L from Sigma
  # GOAL: draw each entry of u_D from standard Gaussian
  u_D = random_state.randn(D) # use random_state.randn(...)
  # GOAL: Want x_D ~ Gaussian(mean = m_D, covar=Sigma_DD)
  x_D = L_DD @ u_D + m_D  # transform u_D into x_D
  return x_D
```

2a: Problem Statement

Show that the Metropolis-Hastings transition distribution $\mathcal{T}$ satisfies detailed balance with respect to the target distribution p^*.

That is, show that:

$$p^*(a)\mathcal{T}(b|a) = p^*(b)\mathcal{T}(a|b) \tag{14}$$

for all possible $a \neq b$, where a, b are any two distinct values of the random variable.

2a: Solution

Substituting the Metropolis-Hastings transition distribution, we have

$$p^*(a)T(b|a) = p^*(a)Q(b|a)\min\left(1, \frac{\tilde{p}(b)Q(a|b)}{\tilde{p}(a)Q(b|a)}\right) \tag{15}$$

$$p^*(b)T(a|b) = p^*(b)Q(a|b)\min\left(1, \frac{\tilde{p}(a)Q(b|a)}{\tilde{p}(b)Q(a|b)}\right) \tag{16}$$

Using the identity, we can match the two expression in 15 and 16.

$$\min\left(1, \frac{x}{y}\right) \cdot y = x \cdot \min\left(1, \frac{y}{x}\right) \tag{17}$$

For this purpose, we substitute $x = \tilde{p}(b)Q(a|b)$ and $y = \tilde{p}(a)Q(b|a)$. Then, we obtain:

$$p^*(a)Q(b|a)\min\left(1, \frac{\tilde{p}(b)Q(a|b)}{\tilde{p}(a)Q(b|a)}\right) = p^*(b)Q(a|b)\min\left(1, \frac{\tilde{p}(a)Q(b|a)}{\tilde{p}(b)Q(a|b)}\right) \tag{18}$$

Expression 18 proves that the detailed balance equation $p^*(a)T(b|a) = p^*(b)T(a|b)$ holds true.

3a: Problem Statement

(See diagram on 3a Instructions web page)

You start at Medford/Tufts station, and take 1000 steps. What is your probability distribution over ending this journey at each of the 7 stations? Report as a vector (use order of nodes in the diagram, small to large). Round to 3 decimal places.

3a: Solution

```
import numpy as np

# Define the transition matrix
P = np.array([
    [0, 1, 0, 0, 0, 0, 0],
    [0.5, 0, 0.5, 0, 0, 0, 0],
    [0, 0.5, 0, 0.5, 0, 0, 0],
    [0, 0, 0.5, 0, 0.5, 0, 0],
    [0, 0, 0, 0.5, 0, 0.5, 0],
    [0, 0, 0, 0, 0.5, 0, 0.5],
    [0, 0, 0, 0, 0, 1, 0]
])

current_state = np.array([1, 0, 0, 0, 0, 0, 0])

for _ in range(1000):
    current_state = np.dot(current_state, P)

print(current_state)
```

[0.167 0.000 0.333 0.000 0.333 0.000 0.167]

3b: Problem Statement

Is there a unique stationary distribution for this Markov chain? If so, explain why. If not, explain why not.

3b: Solution

There is a unique stationary distribution for the Green Line Extension Markov chain because it presents irreducibility and aperiodicity. We see that it is irreducibile since every station can reach any other given enough steps. It is also aperiodic because there are no cycles and there are multiple transition options.