Project 6 Sampling and variational inference

Team K - Minghang Li

April 08, 2023

I'm sorry that I didn't manage to finish the project on time :(but I do want to get feedback for my finished version.

Please give points using my unfinished version – or if you would be so kind, give points using my finished version with late submission deduction for the parts I failed to finish on time.

Problem 15: Monte Carlo estimation of an expected value

Proof that $\mathbb{E}[\hat{g}(\mathbf{X})] = \mathbf{E}[g(X)]$

Proof. It's almost trivial that $\mathbb{E}[g(X)] = \mathbb{E}[g(X_i)]$ (because they are i.i.d from the same probability distribution as X).

$$\begin{aligned}
\mathbb{E}[\hat{g}(\mathbf{X})] &= \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} g(X_i)\right] \\
&= \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N} g(X_i)\right] \\
&= \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[g(X_i)] \\
&= \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[g(X)] \\
&= \frac{N}{N}\mathbb{E}[g(X)] \\
&= \mathbb{E}[g(X)]
\end{aligned}$$

□

Proof that $Var(\hat{g}(\mathbf{X})) = \frac{Var(g(X))}{N}$

Proof. By Bienayme's identity, we know that for pairwise independent variables, we have $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$.

And we also have $Var(g(X_i)) = Var(g(X))$ (by the same reason stated in the first subquestion, they are i.i.d with the same variance as $Var(g(X))$).

The proof then follows:

$$
\begin{aligned}
Var(\hat{g}(X)) &= Var\left(\frac{1}{N}\sum_{i=1}^{N} g(X_i)\right) \\
&= \frac{1}{N^2}Var\left(\sum_{i=1}^{N} g(X_i)\right) \\
&= \frac{1}{N^2}\sum_{i=1}^{N} Var(g(X_i)) \\
&= \frac{1}{N^2}\sum_{i=1}^{N} Var(g(X)) \\
&= \frac{N}{N^2}Var(g(X)) \\
&= \frac{Var(g(X))}{N}
\end{aligned}
$$

$\square$

Problem 16: Sampling in the Rain Network

(a) Derive the expressions

All "$= T$"s are grayed out in the derivation for the ease of my brain to interpret.

Recall that given a Bayesian network with nodes in set X, for a certain node x we have $P(x|X_{\setminus x}) = P(x|MB(x))$.

From the structure of Markove Chain we can know that $S \perp R|C$

With the knowledge learned at hand, we can derive the expressions:

1. Derive $P(C_{=T}|R_{=T}, S_{=T}, W_{=T})$

$$
\begin{aligned}
P(C_{=T}|R_{=T}, S_{=T}, W_{=T}) &= P(C_{=T}|R_{=T}, S_{=T}) && \text{(because } MB(C) = \{R, S\}\text{)} \\
&= \frac{P(R_{=T}, S_{=T}|C_{=T})P(C_{=T})}{P(R_{=T}, S_{=T})} \\
&= \frac{P(R_{=T}|C_{=T})P(S_{=T}|C_{=T})P(C_{=T})}{\sum_C P(R_{=T}|C)P(S_{=T}|C)P(C)} && \text{(because } S \perp R|C\text{)}
\end{aligned}
$$

From the Bayesian network we know that:

$$
\begin{aligned}
P(C_{=T}) &= 0.5 \\
P(R_{=T}|C_{=T}) &= 0.8 \\
P(S_{=T}|C_{=T}) &= 0.1 \\
P(R_{=T}, S_{=T}) &= P(R_{=T}|C_{=T})P(S_{=T}|C_{=T})P(C_{=T}) + P(R_{=T}|C_{=F})P(S_{=T}|C_{=F})P(C_{=F}) \\
&= 0.8 \cdot 0.1 \cdot 0.5 + 0.2 \cdot 0.5 \cdot 0.5 \\
&= 0.09
\end{aligned}
$$

Plug in the values, $P(C_{=T}|R_{=T}, S_{=T}, W_{=T})$ is hence calculated (by the following code blocks) to be $\frac{4}{9}$

```r
P_C_T <- 0.5
P_C_F <- 0.5
P_R_T_given_C_T <- 0.8
P_R_T_given_C_F <- 0.2
P_S_T_given_C_T <- 0.1
P_S_T_given_C_F <- 0.5

P_R_T_and_S_T <- P_R_T_given_C_T * P_S_T_given_C_T * P_C_T +
           P_R_T_given_C_F * P_S_T_given_C_F * P_C_F
```

```r
P_C_T_give_R_T_and_S_T <- (P_R_T_given_C_T * P_S_T_given_C_T * P_C_T) / P_R_T_and_S_T
P_C_T_give_R_T_and_S_T
```

```
## [1] 0.4444444
```

2. Derive $P(C_{=T}|R_{=F},S_{=T},W_{=T})$

$$\begin{aligned} P(C_{=T}|R_{=F},S_{=T},W_{=T}) &= P(C_{=T}|R_{=F},S_{=T}) && \text{(because } MB(C)=\{R,S\}\text{)} \\ &= \frac{P(R_{=F},S_{=T}|C_{=T})P(C_{=T})}{P(R_{=F},S_{=T})} \\ &= \frac{P(R_{=F}|C_{=T})P(S_{=T}|C_{=T})P(C_{=T})}{\sum_C P(R_{=F}|C)P(S_{=T}|C)P(C)} && \text{(because } S \perp R|C\text{)} \end{aligned}$$

We know that S and R are binary variables that only take `True` or `False`, i.e., $P(R_{=T}) + P(R_{=F}) = 1$ and $P(S_{=T}) + P(S_{=F}) = 1$.

Plug in the values, $P(C_{=T}|R_{=F},S_{=T},W_{=T})$ is hence calculated (by the following code blocks) to be $\frac{1}{21}$

```r
P_R_F_given_C_T <- 1 - P_R_T_given_C_T
P_R_F_given_C_F <- 1 - P_R_F_given_C_T
P_R_F_and_S_T <- P_R_F_given_C_T * P_S_T_given_C_T * P_C_T +
           P_R_F_given_C_F * P_S_T_given_C_F * P_C_F
```

```r
P_C_T_give_R_F_and_S_T <- (P_R_F_given_C_T * P_S_T_given_C_T * P_C_T) / P_R_F_and_S_T
P_C_T_give_R_F_and_S_T
```

```
## [1] 0.04761905
```

3. Derive $P(R_{=T}|C_{=T},S_{=T},W_{=T})$

In the following derivation the entities that are canceled out inside the condition probability are either because $S \perp R|C$ or $P(W|X_{\setminus W}) = P(W|MB(W))$.

$$
\begin{aligned}
P(R_{=T}|C_{=T},S_{=T},W_{=T}) &= \frac{P(R_{=T},C_{=T},S_{=T},W_{=T})}{P(C_{=T},S_{=T},W_{=T})} \\
&= \frac{P(W_{=T}|R_{=T},C_{=T},S_{=T})P(R_{=T},C_{=T},S_{=T})}{P(W_{=T}|C_{=T},S_{=T})P(C_{=T},S_{=T})} \\
&= \frac{P(W_{=T}|R_{=T},\cancel{C_{=T}},S_{=T})P(R_{=T}|C_{=T},\cancel{S_{=T}})\cancel{P(C_{=T},S_{=T})}}{P(W_{=T}|C_{=T},S_{=T})\cancel{P(C_{=T},S_{=T})}} \\
&= \frac{P(W_{=T}|R_{=T},S_{=T})P(R_{=T}|C_{=T})}{\sum_R P(W_{=T},R|C_{=T},S_{=T})} \\
&= \frac{P(W_{=T}|R_{=T},S_{=T})P(R_{=T}|C_{=T})}{\sum_R P(W_{=T}|R,\cancel{C_{=T}},S_{=T})P(R|C_{=T},\cancel{S_{=T}})} && \text{(Trying to make use of } P(W|R,S)\text{)} \\
&= \frac{P(W_{=T}|R_{=T},S_{=T})P(R_{=T}|C_{=T})}{\sum_R P(W_{=T}|R,S_{=T})P(R|C_{=T})}
\end{aligned}
$$

From the Bayesian network we know that

$$
\begin{aligned}
P(W_{=T}|R_{=T},S_{=T}) &= 0.99 \\
P(R_{=T}|C_{=T}) &= 0.8 \\
\sum_R P(W_{=T}|R,S_{=T})P(R|C_{=T}) &= P(W_{=T}|R_{=T},S_{=T})P(R_{=T}|C_{=T}) + P(W_{=T}|R_{=F},S_{=T})P(R_{=F}|C_{=T}) \\
&= 0.99 \cdot 0.8 + 0.9 \cdot (1 - 0.8) \\
&= 0.972
\end{aligned}
$$

Plug in the values, we can find that $P(R_{=T}|C_{=T},S_{=T},W_{=T}) = \frac{22}{27}$

```
P_W_T_given_R_T_and_S_T <- 0.99
P_W_T_given_R_F_and_S_T <- 0.9
P_W_T_given_R_T_and_S_F <- 0.9
P_W_T_given_R_F_and_S_F <- 0.1
```

```
P_W_T_given_S_T_and_C_T <- P_W_T_given_R_T_and_S_T * P_R_T_given_C_T +
                           P_W_T_given_R_F_and_S_T * P_R_F_given_C_T
```

```
P_R_T_given_C_T_and_S_T_and_W_T <- (P_W_T_given_R_T_and_S_T * P_R_T_given_C_T) /
                                   P_W_T_given_S_T_and_C_T
P_R_T_given_C_T_and_S_T_and_W_T
```

```
## [1] 0.8148148
```

4. Derive $P(R_{=T}|C_{=F},S_{=T},W_{=T})$

Following similar derivation as $P(R_{=T}|C_{=T},S_{=T},W_{=T})$

$$
\begin{aligned}
&P(R_{=T}|C_{=F},S_{=T},W_{=T}) \\
&\qquad = \frac{P(W_{=T}|R_{=T},S_{=T})P(R_{=T}|C_{=F})}{\sum_R P(W_{=T}|R,S_{=T})P(R|C_{=F})} \\
&\qquad = \frac{0.99 \cdot (1 - 0.8)}{0.99 \cdot (1 - 0.8) + 0.9 \cdot (1 - 0.2)} \\
&\qquad = \frac{11}{51}
\end{aligned}
$$

```
P_R_T_given_C_F_and_S_T_and_W_T <- (P_W_T_given_R_T_and_S_T * P_R_T_given_C_F) /
                          (P_W_T_given_R_T_and_S_T * P_R_T_given_C_F +
                           P_W_T_given_R_F_and_S_T * P_R_F_given_C_F)
P_R_T_given_C_F_and_S_T_and_W_T
```

```
## [1] 0.2156863
```

2. Implement the Gibbs sampler for the Bayesian network

```
## Loading required package: foreach

## Loading required package: iterators

## Loading required package: parallel

## Loading required package: rngtools

## Rlab 4.0 attached.

##
## Attaching package: 'Rlab'

## The following objects are masked from 'package:stats':
##
##    dexp, dgamma, dweibull, pexp, pgamma, pweibull, qexp, qgamma,
##    qweibull, rexp, rgamma, rweibull

## The following object is masked from 'package:datasets':
##
##    precip
```

```
sample_CgivenR <- function(R){
 if (R == F) {
  C <- rbern(1, p=1/21)
 } else {
  C <- rbern(1, p=4/9)
 }
 return(C)
}

sample_RgivenC <- function(C){
  if (C == F) {
   R <- rbern(1, p=11/51)
  } else {
   R <- rbern(1, p=22/27)
  }
 return(R)
}
```

```r
set.seed(2023)
```

```r
gibbs_sampler <- function(niter) {
 C <- rep(0,niter)
 R <- rep(0,niter)
 C[1]=1
 R[1]=1
 foreach (i=2:niter) %dorng% {
  C[i] <- sample_CgivenR(R[i-1])
  R[i] <- sample_RgivenC(C[i])
 }
 res <- data.frame(C=C,R=R)
 print(res)
 res_t <- table(res) / niter
 return(res_t)
}
```

```r
res <- gibbs_sampler(100)
```

```
## Warning: executing %dopar% sequentially: no parallel backend registered
```

```
##    C R
## 1   1 1
## 2   0 0
## 3   0 0
## 4   0 0
## 5   0 0
## 6   0 0
## 7   0 0
## 8   0 0
## 9   0 0
## 10  0 0
## 11  0 0
## 12  0 0
## 13  0 0
## 14  0 0
## 15  0 0
## 16  0 0
## 17  0 0
## 18  0 0
## 19  0 0
## 20  0 0
## 21  0 0
## 22  0 0
## 23  0 0
## 24  0 0
## 25  0 0
## 26  0 0
## 27  0 0
## 28  0 0
## 29  0 0
## 30  0 0
```

```
## 31  0 0
## 32  0 0
## 33  0 0
## 34  0 0
## 35  0 0
## 36  0 0
## 37  0 0
## 38  0 0
## 39  0 0
## 40  0 0
## 41  0 0
## 42  0 0
## 43  0 0
## 44  0 0
## 45  0 0
## 46  0 0
## 47  0 0
## 48  0 0
## 49  0 0
## 50  0 0
## 51  0 0
## 52  0 0
## 53  0 0
## 54  0 0
## 55  0 0
## 56  0 0
## 57  0 0
## 58  0 0
## 59  0 0
## 60  0 0
## 61  0 0
## 62  0 0
## 63  0 0
## 64  0 0
## 65  0 0
## 66  0 0
## 67  0 0
## 68  0 0
## 69  0 0
## 70  0 0
## 71  0 0
## 72  0 0
## 73  0 0
## 74  0 0
## 75  0 0
## 76  0 0
## 77  0 0
## 78  0 0
## 79  0 0
## 80  0 0
## 81  0 0
## 82  0 0
## 83  0 0
## 84  0 0
```

```
## 85  0 0
## 86  0 0
## 87  0 0
## 88  0 0
## 89  0 0
## 90  0 0
## 91  0 0
## 92  0 0
## 93  0 0
## 94  0 0
## 95  0 0
## 96  0 0
## 97  0 0
## 98  0 0
## 99  0 0
## 100 0 0
```

```
res
```

```
##    R
## C      0    1
##   0 0.99 0.00
##   1 0.00 0.01
```