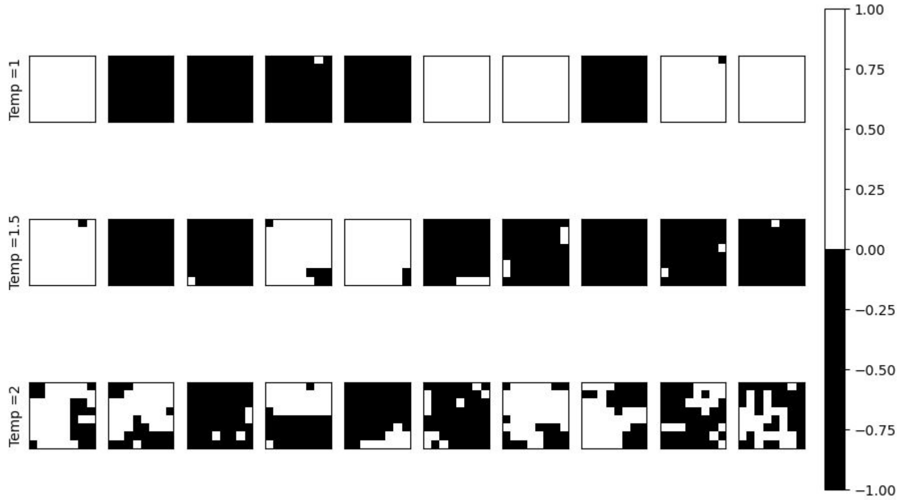
GMDL HW1

Yair Gross, Neta Elmaliach, Sharon Hendy

08/05/2023

Exercise 7

Figure 1: An exact sampler for each of the temperatures: [1, 1.5, 2]



Problem 1

For Temp = 2 in Python 2 we received the result 0.
1/Temp = 0 so Temp$\rightarrow \infty$.
In that case,

$$\lim_{Temp \to \infty} p(x) = \lim_{Temp \to \infty} \exp\left(\frac{1}{Temp} \sum_{s \sim t} (x_s x_t)\right) = \exp(0) = 1$$

so:

$$\lim_{Temp \to \infty} p(x) \propto const$$

Therefore, it means that it is uniform distribution.
In Python 2 the operator / is integer divider when he gets integers values, so when he gets temp = 1.5 it was ok and when he gets the value temp = 2 the result was zero.
(in Python 3 we get the result 0.5).

Exercise 8

draw 10,000 samples and compute two empirical expectations:

$\hat{E}_1(X_{(1,1)}X_{(2,2)}) = 0.9528$
$\hat{E}_1(X_{(1,1)}X_{(8,8)}) = 0.9052$

$\hat{E}_{1.5}(X_{(1,1)}X_{(2,2)}) = 0.7744$
$\hat{E}_{1.5}(X_{(1,1)}X_{(8,8)}) = 0.5658$

$\hat{E}_2(X_{(1,1)}X_{(2,2)}) = 0.5206$
$\hat{E}_2(X_{(1,1)}X_{(8,8)}) = 0.1264$

Problem 2

The closer the points are to each other, the greater the value of E, and vice versa.
This is because there is less chance of disagreement along the way.
The higher the temperature, the more uniformly the probability is distributed - there is a greater chance of disagreement so the value of E decreasing.

Exercise 9

Table 1: Dynamic Programming

Temp	$\hat{E}_{Temp}(X_{(1,1)}X_{(2,2)})$	$\hat{E}_{Temp}(X_{(1,1)}X_{(8,8)})$
1	0.9528	0.9052
1.5	0.7744	0.5658
2	0.5206	0.1264

Table 2: Method 1

Temp	$\hat{E}_{Temp}(X_{(1,1)}X_{(2,2)})$	$\hat{E}_{Temp}(X_{(1,1)}X_{(8,8)})$
1	0.933	0.5338
1.5	0.7556	0.3752
2	0.506	0.071

Table 3: Method 2

Temp	$\hat{E}_{Temp}(X_{(1,1)}X_{(2,2)})$	$\hat{E}_{Temp}(X_{(1,1)}X_{(8,8)})$
1	0.9490	0.9479
1.5	0.7693	0.7217
2	0.5061	0.2429

Problem 3

The difference between the results of the three methods comes from the fact that with Gibbs sampling, for each node the computation of the distribution considers only the values of the node's neighbors in the current sweep. So, the values of the expectations will be less accurate than with Dynamic Programming. It can be seen that the results of $\hat{E}_{Temp}(X_{(1,1)}X_{(2,2)})$ were similar between the two methods compared to the results of $\hat{E}_{Temp}(X_{(1,1)}X_{(8,8)})$, in which there is a greater difference.
The reason for this is that between $X_{(1,1)}$ and $X_{(8,8)}$ there are more sites on the path,, and therefore greater chance for inaccuracy in computing the distribution,

so the difference is greater.
Furthermore, we can see that the difference between the results of method 1 and the DP method are more significant than those between method 2 and the DP method. That could be explained by the fact that in method 2 the expectation is the empirical average of the nodes' values over all the sweeps (apart from the first 100). So, it is less affected by inaccuracies of each individual sweep, as opposed to the calculation in method 1.

Exercise 10

Figure 2: All five images (from the 5 steps- left to right) on a single plot