CSC473W18: Homework Assignment #2: Solutions

Question 1. (10 marks)

a. Notice first that if there is an edge from σ to σ', then there is also one from σ' to σ. In particular, if we can get σ' from σ by swapping σ_i and σ_j, then we can get σ from σ' by swapping σ'_i and σ'_j. Then, to show that G is strongly connected, it suffices to show that there is a path from every permutation σ to the identity permutation $\sigma^* = 1, 2, 3, \ldots, n$. Indeed, we can run a sorting algorithm that works by swapping elements, e.g. Bubble Sort, on σ to sort it into increasing order. Every time Bubble sort swaps two elements, we move along the corresponding edge in the graph G. At the end of the algorithm's execution σ will be sorted, i.e. we will be at σ^*. This describes a path from any permutation σ to σ^*, and because we have edges in both directions between any two adjacent nodes, there is also a path from σ^* to σ. Then, there is also a path from any permutation σ to any other permutation σ': simply follow the path from σ to σ^*, and then the path from σ^* to σ'.

b. The degree of any vertex of G is $\binom{n}{2}$ because we have one edge going out of any σ for any distinct i and and j for which σ_i and σ_j can be swapped. Then, the probability to transition from σ to σ' is

$$p_{\sigma,\sigma'} = \frac{2}{n(n-1)} \min\{1, 2^{\mathrm{inv}(\sigma)-\mathrm{inv}(\sigma')}\}.$$

The probability of staying at σ is

$$p_{\sigma,\sigma} = 1 - \sum_{\sigma' \neq \sigma} p_{\sigma,\sigma'}.$$

c. First we implement a simple procedure to compute $\mathrm{inv}(\sigma)$:

```
Inv(σ)
inv = 0
for i = 1 to n − 1
      for j = i + 1 to n
            if σ_i > σ_j
                  inv = inv + 1
return inv
```

This procedure takes time $O(n^2)$ and uses constant space. Now we implement Metropolis-Hastings as follows. We sample a pair of indices i, j such that $i \neq j$ uniformly at random. Then we define the permutation σ' by $\sigma'_i = \sigma_j$ and $\sigma'_j = \sigma_i$, and $\sigma'_k = \sigma_k$ for all $k \notin \{i, j\}$. We compute $\mathrm{inv}(\sigma)$ and $\mathrm{inv}(\sigma')$ using the procedure above, and if $\mathrm{inv}(\sigma) \geq \mathrm{inv}(\sigma')$ then we transition to σ'. Otherwise, we transition to σ' with probability $2^{\mathrm{inv}(\sigma)-\mathrm{inv}(\sigma')}$, and with the remaining probability we stay at σ.

Question 2. (10 marks)

a. For each j we have

$$\mathbb{E}[c_j^2] = \mathbb{E}[(f_1 h(1) + \ldots + f_n h(n))^2] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} f_i f_j h(i)h(j)\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} f_i f_j \mathbb{E}[h(i)h(j)].$$

For any $i \neq j$, by assumption we have that $\mathbb{E}[h(i)h(j)] = \mathbb{E}[h(i)]\mathbb{E}[h(j)]$, because $h(i)$ and $h(j)$ are independent. But we also assumed that $h(i)$ and $h(j)$ are equally likely to be -1 or $+1$, so we have $\mathbb{E}[h(i)] = \mathbb{E}[h(j)] = 0$. Then, the formula on the right hand side above simplifies to

$$\mathbb{E}[c_j^2] = \sum_{i=1}^{n} f_i^2 \mathbb{E}[h(i)^2] = \sum_{i=1}^{n} f_i^2.$$

Therefore, $\mathbb{E}[c_j^2] = F_2$, the second moment of the stream. It follows that $\mathbb{E}[c] = \frac{1}{k}(\mathbb{E}[c_1^2] + \ldots + \mathbb{E}[c_k^2]) = F_2$ as well.

b. We start with a proof of the inequality in the bonus question. We have

$$\begin{aligned}
\mathbb{E}[c_j^4] &= \mathbb{E}\left[\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} f_{i_1} f_{i_2} f_{i_3} f_{i_4} h(i_1)h(i_2)h(i_3)h(i_4)\right] \\
&= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} f_{i_1} f_{i_2} f_{i_3} f_{i_4} \mathbb{E}[h(i_1)h(i_2)h(i_3)h(i_4)]
\end{aligned}$$

If some i_1 appears in a term in the sum above exactly once, then

$$\mathbb{E}[h(i_1)h(i_2)h(i_3)h(i_4)] = \mathbb{E}[h(i_1)]\mathbb{E}[h(i_2)h(i_3)h(i_4)] = 0.$$

The same logic holds if i_2, i_3, or i_4 appears exactly once. Then, the only terms that are not zero above are those in which every index appears two, or four times. I.e., we have

$$\begin{aligned}
\mathbb{E}[c_j^4] &= \sum_{i=1}^{4} f_i^4 \mathbb{E}[h(i)^4] + 6\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} f_i^2 f_j^2 \mathbb{E}[h(i)^2]\mathbb{E}[h(j)^2] \\
&= \sum_{i=1}^{4} f_i^4 + 6\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} f_i^2 f_j^2 \\
&\leq 3\sum_{i=1}^{4} f_i^4 + 6\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} f_i^2 f_j^2 \\
&= 3\left(\sum_{i=1}^{n} f_i^2\right) = 3\mathbb{E}[c_j^2]^2.
\end{aligned}$$

Then, $\mathrm{Var}(c_j^2) = \mathbb{E}[c_j^4] - \mathbb{E}[c_j^2]^2 \leq 2\mathbb{E}[c_j^2]^2 = 2\mathbb{E}[c]^2$. Finally, because $c_1, \ldots, c_k$ are independent,

$$\mathrm{Var}(c) = \frac{1}{k^2}(\mathrm{Var}(c_1^2) + \ldots \mathrm{Var}(c_k^2)) \leq \frac{2}{k}\mathbb{E}[c]^2.$$

Using Chebyshev's inequality,

$$\mathbb{P}(|c - \mathbb{E}[c]| > \varepsilon \mathbb{E}[c]) \leq \frac{2}{k\epsilon^2}.$$

Therefore, it is enough to set $k \geq \frac{1}{4\epsilon^2}$.

Question 3. (10 marks) The variables of the new linear program (LP) will be the old variables $x_1, \ldots, x_n$ together with new sets of variables $y_{i,1}, \ldots, y_{i,n-3}$, one for each of the m constraints of the original LP. We replace the i-th constraint of the original LP by the $n-3$ constraints

$$\begin{aligned} a_{i,1}x_1 + a_{i,2}x_2 + y_{i,1} &\leq b_i \\ y_{i,1} &= a_{i,3}x_3 + y_{i,2} \\ y_{i,2} &= a_{i,4}x_3 + y_{i,3} \\ &\cdots \\ y_{i,n-3} &= a_{i,n-1}x_{n-1} + a_{i,n}x_n \end{aligned}$$

We can express any equality constraint by two inequalities, for example we can write $y_{i,1} = a_{i,3}x_3 + y_{i,2}$ as

$$\begin{aligned} y_{i,1} - a_{i,3}x_3 - y_{i,2} &\leq 0 \\ a_{i,3}x_3 + y_{i,2} - y_{i,1} &\leq 0 \end{aligned}$$

It is easy to show by induction on $n-2-j$ that the constraints ensure that $y_{i,j} = a_{i,j+2}x_{j+2} + \ldots + a_{i,n}x_n$. Then, the first constraint is equivalent to the original constraint $a_{i,1}x_1 + \ldots a_{i,n}x_n \leq b_i$. The objective function of the new LP is the same as that of the old LP: maximize $c_1x_1 + \ldots + c_nx_n$. The new variables have coefficient zero in the objective function. Becaue the objective function is unchanged, and the original variables satisfy the same constraints as in the original LP, the two LPs have the same optimal objective value. Moreover, taking an optimal solution of the new LP and restricting its variables to the original variables $x_1, \ldots, x_n$ gives an optimal solution to the original LP.

It is clear that this construction can be done in polynomial time: we merely have to replace each constraint of the original LP by the $n-3$ new constraints, which can be done in time $O(mn)$.