Yale University
CPSC 516, Spring 2023
Assignment 4

Chang Feng (Felix) Zhou cz397

P.1.

(a)

For the (i, j)-th entry of A, let us write it as

$$A_{ij} = \frac{a_{ij}}{b_{ij}}$$

for some coprime $a_{ij} \in \mathbb{Z}, b_{ij} \in \mathbb{Z}_+$. Then consider

$$M := \prod_{i,j} b_{ij} \in \mathbb{Z}$$
$$B := MA.$$

By construction, $A = \frac{1}{M}B$ and since $b_{ij} \mid M$ for all i, j, we know that $A \in \mathbb{Z}^{m\times n}$. The bit complexity of M is

$$O\left(\log \prod_{i,j} b_{ij}\right) = O\left(\sum_{i,j} \log b_{ij}\right).$$

This is at most the bit complexity of A. The bit complexity of any B_{ij} is at most

$$O\left(\log a_{ij} \prod_{k,\ell} b_{k,\ell}\right) = O\left(\log a_{ij} + \sum_{k,\ell} \log b_{k,\ell}\right).$$

Once again, this is at most the bit complexity of A.

(b)

Suppose $C \in \mathbb{Q}^{p\times p}$. Write

$$D := MC \in \mathbb{Z}^{p\times p}.$$

The matrix norm is obtained by some unit-vector $x \in \mathbb{R}^p$. Thus

$$
\begin{aligned}
\|C\|_2 &= \|Cx\| \\
&= \sqrt{\sum_{i=1}^{p} (Cx)_i^2} \\
&\leq \sum_{i=1}^{p} |(Cx)_i| && \sqrt{a+b} \leq \sqrt{a} + \sqrt{b} \\
&\leq \sum_{i=1}^{p} \|C^{(i)}\| \cdot \|x\| && C^{(i)} \ i\text{-th row} \\
&= \sum_{i=1}^{p} \|C^{(i)}\| \\
&\leq \sum_{i=1}^{p} \sum_{j=1}^{p} |C_{ij}| \\
&= \frac{1}{M} \sum_{i=1}^{p} \sum_{j=1}^{p} |D_{ij}|.
\end{aligned}
$$

The bit complexity of D_{ij} is at most L. Hence the value of the summation is at most $n^2 2^L$. Thus the operator norm is at most

$$2^{O(L \log n)}$$

as desired.

On the other hand,

$$\|C^{-1}\| = \frac{M}{|\det D|} \|\mathrm{adj}(D)\|.$$

The equality follows by Cramer's rule.

By cofactor expansion, $\det D \in \mathbb{Z}$, thus we may as well ignore it. M has bit complexity at most L and hence contributes at most a multiplicative 2^L factor to the matrix norm.

Now, the entries of the adjugate matrix are determinants of submatrices of D. In particular, it is the product of singular values and is thus at most the largest singular value of D to the n-th power. But the largest singular value is simply the operator norm of D, which is at most the sum of absolute values of entries in D from our work above. Each entry of D has bit complexity at most L. Thus the largest singular value is at most $n^2 2^L$. Taking this to the power of n and accounting for M yields the bound

$$2^{O(nL \log n)}.$$

(c)

Let x be a vertex of K and recall that it is uniquely determined by some invertible submatrix C of A so that $Cx = b^=$ where $b^=$ is some subvector of b.

Write $b^= = b'/B$ where B is the product of denominators in $b^=$. Using the notation from P.1.(b), we have

$$x = C^{-1} b^= = \frac{M}{B \det C} \mathrm{adj}(C) b'.$$

But since all intermediaries are rational, x must be rational as well.

We showed in P.1.(b) (indirectly) that the bit complexity of $M \operatorname{adj} C$ is $O(nL \log n)$. Now,

$$\begin{aligned}(M \operatorname{adj}(C) b')_i &= M \sum_j \operatorname{adj}(C)_{ij} b'_j \\ &\le M 2^L \sum_j \operatorname{adj}(C)_{ij}.\end{aligned}$$

The inequality comes from the fact that b'_j has bit complexity at most L. It follows that $M \operatorname{adj}(C) b'$ has bit complexity at most $O(nL \log n)$ since $M \operatorname{adj} C$ has bit complexity $O(nL \log n)$.

The bit complexity of $B, \det C$ are all at most $O(nL \log n)$. Hence x has bit complexity at most

$$O(nL \log n)$$

as desired.

P.2.

(a)

Lemma 1:
If $f : J \subseteq \mathbb{R} \to \mathbb{R}, g : K \subseteq \mathbb{R}^m \to J$ are convex and f is non-decreasing, then $f \circ g$ is convex.

Proof : Lemma 1
Fix $\lambda \in [0, 1]$ and $x, y \in K$.

$$\begin{aligned} f(g[(1-\lambda)x + \lambda y]) &\leq f([1-\lambda]g(x) + \lambda g(y)) && f \text{ non-decreasing, } g \text{ convex} \\ &\leq (1-\lambda)f(g(x)) + \lambda f(g(y)). && f \text{ convex} \end{aligned}$$

$\square$

Consider the function $\langle x, \mathbb{1}_M \rangle$. It is a linear function and is thus convex. Now, the exponential is convex and non-decreasing, so $\exp(\langle x, \mathbb{1}_M \rangle)$ is convex. Moreover, the sum of convex function is convex. So $\sum_M \exp(\langle x, \mathbb{1}_M \rangle)$ is convex. Finally, ln is non-decreasing and thus

$$f(x) = \ln \sum_M \exp(\langle x, \mathbb{1}_M \rangle)$$

is convex as desired.

(b)

Let us assume that G has bipartition $V = (U, W)$ where $|U| = |W| = n/2$, or else $P = \varnothing$ and the problem is trivial.

We claim that P is equivalent to the following polytope Q

$$\begin{aligned} \sum_{v \sim u} x_{uv} &= 1 && \forall u \in V \\ x &\geq 0. \end{aligned}$$

If we let A be the vertex-edge incident matrix of G, then we can succinctly write this as

$$\begin{aligned} Ax &= \mathbb{1} \\ x &\geq 0. \end{aligned}$$

Note that for $x \in \{0, 1\}^m$, $x \in Q$ if and only if $x = \mathbb{1}_M$ for some perfect matching M.

If we show $P = Q$ then we are done, since we can just check in polynomial time whether any of the constraints are violated.

To see the claim, first note that $P \subseteq Q$. This is because any indicator vector for a matching $\mathbb{1}_M$ necessarily satisfies all the inequalities. But then all the convex combinations of indicator variables also satisfies the inequalities as well since the inequalities are linear.

It remains to show that $Q \subseteq P$. We argue that the extreme points of Q are integral, ie the extreme points of Q are indicator vectors of perfect matchings. This would complete the proof since Q is then the convex hull of some indicator vectors while P is the convex hull of all indicator vectors.

Lemma 2:
A is totally unimodular, ie every square submatrix of A has determinant taking values in $\{-1, 0, 1\}$.

Proof : Lemma 2

Without loss of generality, assume that G has bipartition $V = (U, W)$ and the rows of A are such that the first $n/2$ correspond to U and the last $n/2$ correspond to W.

Let $B \in \{-1, 0, 1\}^{k \times k}$ be a square submatrix of A. We argue by induction on k.

The base case of $k = 1$ certainly holds.

Suppose inductively that this holds up to $k - 1$. If B has any zero columns, then $\det(B)$ is zero and we are done. Otherwise, if B has any columns with a single non-zero entry 1, we can use cofactor expansion along that column to determine that

$$\det(B) = \pm \det(B')$$

where B' is a $(k-1) \times (k-1)$ submatrix of A. In this case, we are also done. Finally, suppose every column of B has exactly two non-zero entries. But since we assumed that A has the particular format, if we subtract the rows of B corresponding to U from the rows of B corresponding to W, we get the zero vector and thus B is singular.

By induction, we conclude the proof. $\square$

To see why the lemma completes the proof, note that any extreme point x of Q is determined by some invertible square submatrix $A_=$ where the non-zero entries of x are given by

$$x_= = A_=^{-1} \mathbb{1}_= = \frac{1}{\det(A_=)} \operatorname{adj}(A_=) \mathbb{1}_=.$$

But $\frac{1}{\det(A_=)} \in \{\pm 1\}$ and $\operatorname{adj}(A_=)$ is also integral, Hence $x_=$ must be integral as well.

(c)

Suppose we can evaluate $f(\mathbb{1})$ in polynomial time. Then

$$\begin{aligned}
\exp f(\mathbb{1}) &= \sum_{M \in \mathcal{M}} \exp(\langle \mathbb{1}, \mathbb{1}_M \rangle) \\
&= \sum_{M \in \mathcal{M}} \exp(n/2) \\
&= |\mathcal{M}| \exp(n/2) \\
|\mathcal{M}| &= \frac{\exp f(\mathbb{1})}{\exp(n/2)}.
\end{aligned}$$

Thus we can count the number of perfect matching within G in polynomial time.

P.3.

By computation,

$$\begin{aligned}
\nabla f(x) &= \sum_S \frac{\exp\langle x, \mathbb{1}_S\rangle}{\sum_T \exp\langle x, \mathbb{1}_T\rangle}\mathbb{1}_S \\
\nabla_i f(x) &= \sum_{S \ni i} \frac{\exp\langle x, \mathbb{1}_S\rangle}{\sum_T \exp\langle x, \mathbb{1}_T\rangle} \\
\frac{\partial \nabla_i f(x)}{\partial x_j} &= \sum_{S \ni i,j} \frac{\exp\langle x, \mathbb{1}_S\rangle}{\sum_T \exp\langle x, \mathbb{1}_T\rangle} - \nabla_i f(x) \nabla_j f(x).
\end{aligned}$$

Note that when we take a quadratic form, the second term in $\nabla^2 f(x)$ is non-positive.

Lemma 3:
If f is convex, then ∇f is L-Lipschitz if

$$y^T \nabla^2 f(x) y \leq L\|y\|^2$$

for all $y \in \mathbb{R}^n$.

Proof : Lemma 3

Recall we know that $\nabla^2 f \succeq 0$ from the convexity of f. The described condition is equal to the condition that the Rayleigh quotient is at most L. This is equivalent to the largest eigenvalue being at most L. Finally, this is equivalent to the condition that the operator norm of $\nabla^2 f$ is always at most L,

Define $g(t) := \nabla f([1-\lambda]x + \lambda y)$ and remark that $g'(t) = \nabla^2 f([1-\lambda]x + \lambda y)[y - x]$. We can thus leverage the integral representation of g.

$$\begin{aligned}
\|\nabla f(y) - \nabla f(x)\| &= \|g(1) - g(0)\| \\
&= \left\| \int_0^1 g'(t) dt \right\| \\
&\leq \int_0^1 \|g'(t)\| dt \\
&= \int_0^1 \|\nabla^2 f([1-\lambda]x + \lambda y)[y - x]\| dt \\
&\leq \int_0^1 L\|y - x\| dt \\
&= L\|y - x\|.
\end{aligned}$$

□

By the lemma, is suffices to bound $y^T \boldsymbol{\nabla}^2 f(x) y$.

$$\begin{aligned}
\sum_{i,j} y_i y_j \boldsymbol{\nabla}^2_{i,j} f(x) &= \sum_{i,j} y_i y_j \sum_{S \ni i,j} \frac{\exp\langle x, \mathbb{1}_S \rangle}{\sum_T \exp\langle x, \mathbb{1}_T \rangle} - y^T \boldsymbol{\nabla} f(x) \boldsymbol{\nabla} f(x)^T y \\
&\leq \sum_{i,j} y_i y_j \sum_{S \ni i,j} \frac{\exp\langle x, \mathbb{1}_S \rangle}{\sum_T \exp\langle x, \mathbb{1}_T \rangle} \\
&\leq \sum_{i,j} y_i y_j \cdot 1 \\
&\leq n^2 \|y\|_\infty^2 \\
&\leq n^2 \|y\|_2^2.
\end{aligned}$$

By the lemma, $\boldsymbol{\nabla} f$ is n^2-Lipschitz as desired.