

# Exercise of Chapter 8

Xiaoyu Chen

## 1 Exercise 8.2.

Prove Jensen's inequality.

Recall The Jensen's inequality as follow:

Let  $K \subset \mathbb{R}^k$  be a compact convex set,  $X$  a r.v. taking values in  $K$ , and  $g : K \rightarrow \mathbb{R}$  a convex real-valued function. Then  $g(\mathbb{E}X) \leq \mathbb{E}g(X)$

### 1.1 solution

Following the hint, we first construct the graph

$$G(g) = \{(x, y) : x \in K \text{ and } y \geq g(x)\} \subset \mathbb{R}^{k+1}$$

**Claim 1.**  $G(g)$  is a convex set

*Proof.* For any two vertex  $x = (a, b), y = (c, d) \in G(g)$ , and some  $\lambda \in [0, 1]$ , we have

$$\begin{aligned}\lambda x + (1 - \lambda)y &= \lambda(a, b) + (1 - \lambda)(c, d) \\ &= (\lambda a + (1 - \lambda)c, \lambda b + (1 - \lambda)d)\end{aligned}$$

Since  $a \in K$  and  $c \in K$ , and  $K$  is a convex set, we have  $\lambda a + (1 - \lambda)c \in K$ . Moreover, we have:

$$\begin{aligned}\lambda b + (1 - \lambda)d &\geq \lambda g(a) + (1 - \lambda)g(c) \\ &\geq g(\lambda a + (1 - \lambda)c), \quad \text{since } g \text{ is convex}\end{aligned}$$

Combining these together we could find that  $\lambda x + (1 - \lambda)y \in G(g)$ . Then, by the definition of convex set,  $G(g)$  is a convex set.  $\square$

**Claim 2.** we have a supporting hyperplane for  $G(g)$  at  $(\mathbb{E}x, g(\mathbb{E}x))$

*Proof.* Since  $G(g)$  is a convex set, and  $(\mathbb{E}x, g(\mathbb{E}x))$  is on the boundary of  $G(g)$ , there is a supporting hyperplane for  $G(g)$ . (this is followed by the [supporting hyperplane theorem](#))  $\square$

For convenience, we denote this hyperplane by

$$P : a^T(x - b) = 0, \quad \text{where } b = (\mathbb{E}X, g(\mathbb{E}X))$$

For all  $v = (x, y) \in P$  and  $x \in K$  we have  $y \leq g(x)$ , since  $P$  is the supporting hyperplane of  $G(g)$ . This is the key observation of the whole solution.

**Claim 3.** *For any vertex  $a$ , let  $a_{n-1} \in \mathbb{R}^{k-1}$  be the sub-vertex of  $a$ , and  $a_n$  be the last dimension of  $a$ . Then for a hyperplane  $P : a^T(X - b) = 0$  and a point  $x \in P$ , we have:*

$$x_n = \frac{1}{a_n}(a_{n-1}^T(b_{n-1} - x_{n-1})) + b_n$$

*Proof.*

$$\begin{aligned} a^T(x - b) &= 0 \\ a^T x &= a^T b \\ a_{n-1}^T x_{n-1} + a_n x_n &= a_{n-1}^T b_{n-1} + a_n b_n \\ a_n x_n &= a_{n-1}^T b_{n-1} - a_{n-1}^T x_{n-1} + a_n b_n \\ x_n &= \frac{1}{a_n}(a_{n-1}^T(b_{n-1} - x_{n-1})) + b_n \quad \square \end{aligned}$$

So, If the hyperplane  $P$  is the one at  $(\mathbb{E}X, g(\mathbb{E}X))$ , we have

$$f(x) = \frac{1}{a_n}(a_{n-1}^T(\mathbb{E}X - x)) + g(\mathbb{E}X), \quad \text{where } (x, f(x)) \in P \text{ and } x \in K$$

**Claim 4.**  $\mathbb{E}f(Y) = g(\mathbb{E}X)$

*Proof.*

$$\begin{aligned} \mathbb{E}f(X) &= \frac{\int_K f(x) dx}{\int_K dx} \\ &= \frac{\int_K [\frac{1}{a_n}(a_{n-1}^T(\mathbb{E}X - x)) + g(\mathbb{E}X)] dx}{\text{Vol}(K)} \\ &= \frac{\frac{a_{n-1}^T}{a_n} \int_K (\mathbb{E}X - x) dx + g(\mathbb{E}X) \text{Vol}(K)}{\text{Vol}(K)} \\ &= g(\mathbb{E}X) \quad \square \end{aligned}$$

So we have  $g(\mathbb{E}X) = \mathbb{E}f(X) \leq \mathbb{E}g(X)$ , since  $f(x) \leq g(x)$  on every  $x \in K$ .

## 2 Exercise 8.3

Verify (8.1) and (8.2). (One of these identities is actually trivial)

## 2.1 verify (8.2)

Lets start from the trivial one.

$$\mathcal{E}_P(\varphi, \varphi) = \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) P(x, y) (\varphi(x) - \varphi(y))^2$$

Let

$$\begin{aligned} G[\Omega] &= \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) P(x, y) (\varphi(x) - \varphi(y))^2 \\ G[\Omega_1, \Omega_2] &= \sum_{x \in \Omega_1, y \in \Omega_2} \pi(x) P(x, y) (\varphi(x) - \varphi(y))^2 \end{aligned}$$

Then we have,

$$\begin{aligned} \pi(\Omega_0) \mathcal{E}_{P_0}(\varphi, \varphi) + \pi(\Omega_1) \mathcal{E}_{P_1}(\varphi, \varphi) + C \\ = G[\Omega_0] + G[\Omega_1] + G[\Omega_0, \Omega_1] \end{aligned}$$

Since in this definition, we have assumed time reversibility of  $(\Omega, p)$ . So,  $G[\Omega_0]$  enumerates all unordered pairs  $\{x, y\}$  where  $x, y \in \Omega_0$ . Similarly,  $G[\Omega_1]$  enumerates all unordered pairs  $\{x, y\}$  where  $x, y \in \Omega_1$ . Moreover,  $G[\Omega_0, \Omega_1]$  enumerate all the unordered pairs  $\{x, y\}$  where  $x \in \Omega_0$  and  $y \in \Omega_1$ . And hence

$$\pi(\Omega_0) \mathcal{E}_{P_0}(\varphi, \varphi) + \pi(\Omega_1) \mathcal{E}_{P_1}(\varphi, \varphi) + C$$

enumerates all the unordered pairs  $\{x, y\}$  where  $x, y \in \Omega$ , which is the same as  $\mathcal{E}_P(\varphi, \varphi)$ .

## 2.2 verify (8.1)

First, for convenience, lets define some notations.

$$\begin{aligned} [\mathbb{E}\varphi]_{\Omega}^{\pi} &:= \sum_{x \in \Omega} \pi(x) \varphi(x) \\ \varphi &:= \varphi(x), \quad \text{this is an abbreviate} \end{aligned}$$

First, we know that

$$\begin{aligned} \text{Var}_{\pi} \varphi &= \mathbb{E} \varphi^2 - (\mathbb{E} \varphi)^2 \\ \text{Var}_{\pi_0} \varphi &= [\mathbb{E} \varphi^2]_{\Omega_0}^{\pi_0} - ([\mathbb{E} \varphi]_{\Omega_0}^{\pi_0})^2 \\ \text{Var}_{\pi_1} \varphi &= [\mathbb{E} \varphi^2]_{\Omega_1}^{\pi_1} - ([\mathbb{E} \varphi]_{\Omega_1}^{\pi_1})^2 \end{aligned}$$

And we could rewrite  $(\mathbb{E} \varphi)^2$  as

$$\begin{aligned} (\mathbb{E} \varphi)^2 &= (\mathbb{E} \varphi)(\mathbb{E} \varphi) \\ &= [\mathbb{E} \varphi]_{\Omega_a}^{\pi} \mathbb{E} \varphi + [\mathbb{E} \varphi]_{\Omega_b}^{\pi} \mathbb{E} \varphi \\ &= ([\mathbb{E} \varphi]_{\Omega_a}^{\pi})^2 + 2[\mathbb{E} \varphi]_{\Omega_a}^{\pi} [\mathbb{E} \varphi]_{\Omega_b}^{\pi} + ([\mathbb{E} \varphi]_{\Omega_b}^{\pi})^2 \end{aligned}$$

Then

$$\begin{aligned}
& \text{Var}_{\pi} \varphi - \pi(\Omega_0) \text{Var}_{\pi_0} \varphi - \pi(\Omega_1) \text{Var}_{\pi_1} \\
&= \pi(\Omega_0) ([\mathbb{E} \varphi]_{\Omega_0}^{\pi_0})^2 + \pi(\Omega_1) ([\mathbb{E} \varphi]_{\Omega_1}^{\pi_1})^2 - (\mathbb{E} \varphi)^2 \\
&= \frac{1 - \pi(\Omega_0)}{\pi(\Omega_0)} ([\mathbb{E} \varphi]_{\Omega_0}^{\pi})^2 + \frac{1 - \pi(\Omega_1)}{\pi(\Omega_1)} ([\mathbb{E} \varphi]_{\Omega_1}^{\pi})^2 - 2 [\mathbb{E} \varphi]_{\Omega_0}^{\pi} [\mathbb{E} \varphi]_{\Omega_1}^{\pi} \\
&= \frac{\pi(\Omega_1)}{\pi(\Omega_0)} ([\mathbb{E} \varphi]_{\Omega_0}^{\pi})^2 + \frac{\pi(\Omega_0)}{\pi(\Omega_1)} ([\mathbb{E} \varphi]_{\Omega_1}^{\pi})^2 - 2 [\mathbb{E} \varphi]_{\Omega_0}^{\pi} [\mathbb{E} \varphi]_{\Omega_1}^{\pi}, \quad \text{since } \pi(\Omega_0) + \pi(\Omega_1) = 1 \\
&= \pi(\Omega_0) \pi(\Omega_0) ([\mathbb{E} \varphi]_{\Omega_0}^{\pi_0})^2 + \pi(\Omega_0) \pi(\Omega_1) ([\mathbb{E} \varphi]_{\Omega_1}^{\pi_1})^2 - 2 \pi(\Omega_0) \pi(\Omega_1) [\mathbb{E} \varphi]_{\Omega_0}^{\pi_0} [\mathbb{E} \varphi]_{\Omega_1}^{\pi_1} \\
&= \pi(\Omega_0) \pi(\Omega_0) \{ ([\mathbb{E} \varphi]_{\Omega_0}^{\pi_0})^2 + ([\mathbb{E} \varphi]_{\Omega_1}^{\pi_1})^2 - 2 [\mathbb{E} \varphi]_{\Omega_0}^{\pi_0} [\mathbb{E} \varphi]_{\Omega_1}^{\pi_1} \} \\
&= \pi(\Omega_0) \pi(\Omega_0) (\mathbb{E}_{\pi_a} \varphi - \mathbb{E}_{\pi_b} \varphi)^2 \\
&= \text{Var}_{\pi} \bar{\varphi}
\end{aligned}$$

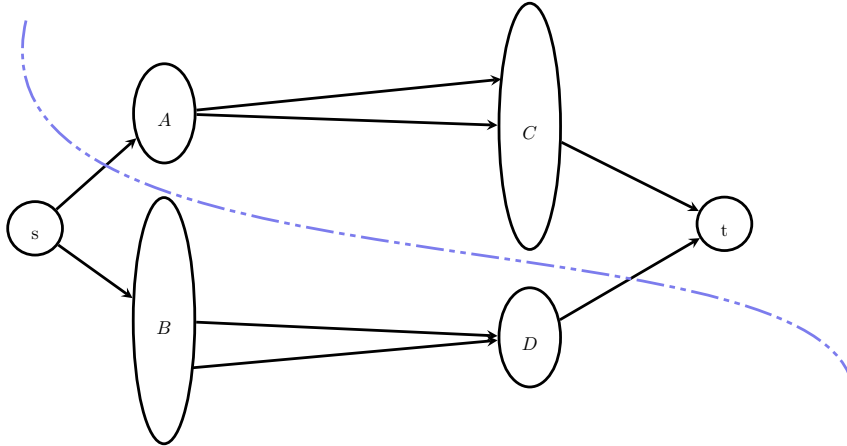
### 3 Exercise 8.13

Prove Lemma 8.12.

#### 3.1 solution

*Proof.* For convenience, let  $\delta(A, B)$  be the cut between  $A$  and  $B$ . Let  $\Omega_0 = \mathcal{B}_{\bar{e}} = A \cup B$  and  $\Omega_1 = \mathcal{B}_e = C \cap D$ . Let  $\pi_0 = \pi/\pi(\Omega_0)$  and  $\pi_1 = \pi/\pi(\Omega_1)$ . Since  $\delta(s, \Omega \setminus \{s\}) = 1$ ,

So, we only need to prove that there does not exist a cut which is less than 1. Thus this cut should not contains any unbounded edge in the middle. This case is showed in the picture below.



Let  $\delta^* = \delta(\{s\} \cup B \cup D, \{t\} \cup A \cup C)$ . If  $\delta^* < 1$ , then

$$\begin{aligned}\pi_0(A) + \pi_1(D) &< 1 \\ \pi_1(D) &< 1 - \pi_0(A) \\ \pi_1(D) &< \pi_0(B)\end{aligned}$$

And we have  $D = \Gamma_e(B)$ , since otherwise,  $\Gamma_e(B) \cap C \neq \emptyset$ , and the cut is unbounded. So

$$\begin{aligned}\pi_1(\Gamma_e(B)) &< \pi_0(B) \\ \frac{|\Gamma_e(B)|}{|\mathcal{B}_e|} &< \frac{|B|}{|\mathcal{B}_e|}, \quad B \subseteq \mathcal{B}_e\end{aligned}$$

But this contradicts with Lemma 8.8, so we could draw a conclusion that  $\delta^* \geq 1$ . Which means the minimum cut in this graph is 1. By applying the Max-flow, min-cut Theorem, we know that this graph has a flow of value 1. Hence, the properties asked by Lemma 8.12 is obvious.  $\square$

## 4 Exercises 8.17

Proving that class of orientable matroids is the same as the class of regular matroids requires familiarity with matroid theory. However, the weaker claim that the cycle matroid of any graph is orientable is an exercise in straight combinatorics. Prove the claim.

### 4.1 solution

Intuitively, as in Figure 1, if a cut intersects  $D$  with a cycle  $C$ , then the number of green edges is equal to the number of red edges. Using this phenomenon, we could construct proper  $\gamma$  and  $\delta$ .

Formally, consider a graphic matroid  $\mathcal{B}$  of a graph  $G$ . Let  $\mathcal{C}$  and  $\mathcal{D}$  be the set of all the cycles and the set of all the cuts in  $\mathcal{B}$ . First, give each edge  $e = \{u, v\} \in G$  a fixed direction  $O_G(e)$ . Then we could define  $\gamma$  and  $\delta$  based on this direction.

For each cycle  $C \in \mathcal{C}$ , we give edge  $e \in C$  another direction  $O_C(e)$  which is determined by an arbitrary traveling through  $C$ . Then:

$$\gamma(C, g) := \begin{cases} 0, & \text{if } g \notin C \\ 1, & \text{if } g \in C \text{ and } O_C(g) = O_G(g) \\ -1, & \text{if } g \in C \text{ and } O_C(g) \neq O_G(g) \end{cases} \quad (1)$$

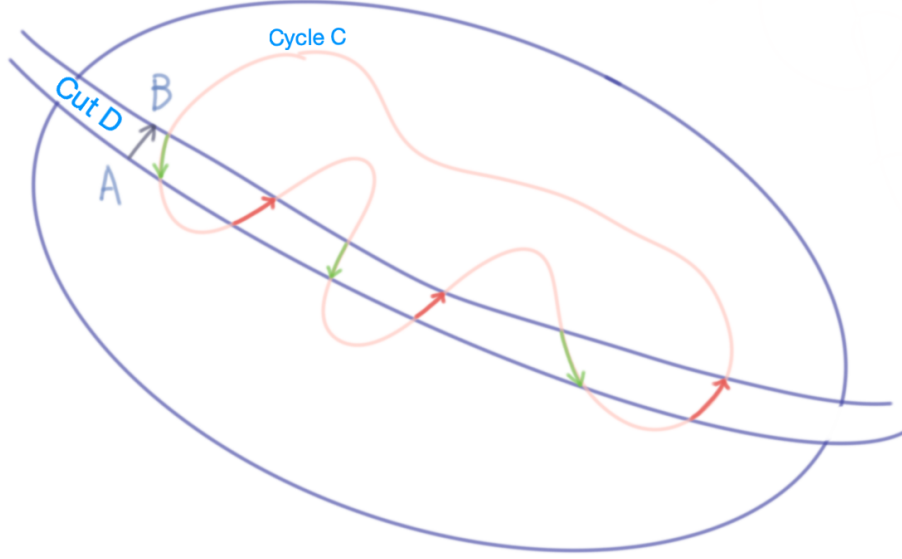


Figure 1: A cut intersects with a cycle. If we make all the edges of the cycle in counterclockwise direction. Then, it is easy for us to find that the number of edges from B to A (the green edges) is equal to the number of edges from A to B (the red edges).

For each cut  $D \in \mathcal{D}$  where  $G \setminus D = (A, B)$ , we give edge  $e \in D$  another direction  $O_D(e)$  which is determined by the order of A and B. Then:

$$\delta(D, g) := \begin{cases} 0, & \text{if } g \notin D \\ 1, & \text{if } g \in C \text{ and } O_D(g) = O_G(g) \\ -1, & \text{if } g \in C \text{ and } O_D(g) \neq O_G(g) \end{cases} \quad (2)$$

Then for any cycle C, cut D and an edge g, we have

$$\gamma(C, g)\delta(D, g) = \begin{cases} 0, & \text{if } g \text{ is not in } C \text{ or } D \\ 1, & \text{if } g \text{ is in } C \text{ and } D \text{ and } O_C(g) = O_D(g) \Rightarrow \text{red edge} \\ -1, & \text{if } g \text{ is in } C \text{ and } D \text{ and } O_C(g) \neq O_D(g) \Rightarrow \text{green edge} \end{cases}$$

Since the number of green edges and the number of red edges are equal, we have

$$\sum_{g \in E} \gamma(C, g)\delta(D, g) = 0 \quad (3)$$

Together with Equation (1) (2) (3), we know that  $\mathcal{B}$  is an orientable matroid.

## 5 Exercies 8.18

Another way to demonstrate that all graphic matroids are balanced is via the theory of electrical networks. Regard a graph  $G = (V, E)$  as an electrical network, with vertices as terminals and edges as unit resistors. The key facts are: (1) For any edge  $e = u, v$ , the effective between vertices  $u$  and  $v$  is equal to  $\tau(G/e)/\tau(G)$ , where  $\tau(G)$  is the number of spanning trees in  $G$ , and  $\tau(G/e)$  is the number of spanning trees in  $G$  that include the edge  $e$ . This result is essentially due to Kirchhoff; see Van Lint and Wilson [79, Thm. 34.3]. (2) If the resistance of some edge of a network is decreased, the effective resistance between any two terminals does not increase. This is “Rayleigh’s Monotonicity Principle”; see Doyle and Snell [25].

### 5.1 solution

Let  $\mathcal{B}$  be the graphic matroid of  $G$ . The effective resistance of  $e$  in  $G$  is  $\frac{|\mathcal{B}_e(M)|}{|\mathcal{B}(M)|}$ . Consider an edge  $f \in E$ , changing the resistance of  $f = \{u, v\}$  from 1 to 0 is equivalent to shrink  $f$  and make  $u, v$  the same point in the network. And if we do so, the effective resistance of  $e$  in the new graph is  $\frac{|\mathcal{B}_e(M/f)|}{|\mathcal{B}(M/f)|}$ . And we know that

$$\begin{aligned} \frac{|\mathcal{B}_e(M/f)|}{|\mathcal{B}(M/f)|} &\leq \frac{|\mathcal{B}_e(M)|}{|\mathcal{B}(M)|} \\ \frac{|\mathcal{B}_{ef}(M)|}{|\mathcal{B}_f(M)|} &\leq \frac{|\mathcal{B}_e(M)|}{|\mathcal{B}(M)|} \\ \Pr(e|f) &\leq \Pr(e) \end{aligned}$$

So, graphic matroid is balanced.