Spectral Graph Theory

R. Kyng & M. Probst Gutenberg *Problem Set 6— Monday, April 8th*

The exercises for this week will not count toward your grade, but you are highly encouraged to solve them all. This exercise sheet has exercises related to week 6. We encourage you to start early so you have time to go through everything.

To get feedback, you must hand in your solutions by 23:59 on April 18. Both hand-written and LaTeX solutions are acceptable, but we will only attempt to read legible text.

Notation

Througout the following exercises, we will use the following notation:

- S^n is the set of symmetric real $n \times n$ matrices.
- S^n_+ is the set of positive semi-definite $n \times n$ matrices.
- S^n_{++} is the set of positive definite $n \times n$ matrices.

Whenever we say a matrix is positive semi-definite or positive definite, we require it to be real and symmetric.

Exercise 1

Consider $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^n$.

1. Assume that $\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top$ is invertible. Determine c such that

$$\left(\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{I} - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{c}.$$

2. Assume that both $\boldsymbol{A}$ and $\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top$ are invertible. Prove that

$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}}.$$

Hint: You might use that $(\boldsymbol{B}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}$ *for two invertible matrices* $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n \times n}$.

Solution.

1. We want to find c such that

$$\boldsymbol{I} = \left(\boldsymbol{I} - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{c}\right)\left(\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top\right).$$

This is equivalent to

$$\boldsymbol{I} = \boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{c} - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{c}\boldsymbol{u}\boldsymbol{v}^\top.$$

Notice that $\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{u}\boldsymbol{v}^\top = \boldsymbol{u}(\boldsymbol{v}^\top\boldsymbol{u})\boldsymbol{v}^\top = (\boldsymbol{v}^\top\boldsymbol{u})\boldsymbol{u}\boldsymbol{v}^\top$. Thus, the above equation holds if and only if

$$1 - \frac{1}{c} - \frac{\boldsymbol{v}^\top\boldsymbol{u}}{c} = 0.$$

Therefore, we set $c = 1 + \boldsymbol{v}^\top\boldsymbol{u}$ and conclude hat

$$\left(\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{I} - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{1 + \boldsymbol{v}^\top\boldsymbol{u}}. \tag{1}$$

2. We have that

$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \left(\boldsymbol{A}\left(\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\right)\right)^{-1}.$$

By applying the fact that $(\boldsymbol{B}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}$ for two invertible matrices $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n\times n}$ and then using Equation (1), we get

$$\begin{aligned}
\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} &= \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\right)^{-1}\boldsymbol{A}^{-1} \\
&= \left(\boldsymbol{I} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top}{1 + \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}}\right)\boldsymbol{A}^{-1} \\
&= \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}}.
\end{aligned}$$

This equality is known as the *Sherman-Morrison* formula.

Exercise 2

Consider a matrix function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$. For $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{n\times n}$, we define

$$Df\left(\boldsymbol{X}\right)\left[\boldsymbol{Y}\right] = \frac{\partial}{\partial t}\bigg|_{t=0} f\left(\boldsymbol{X} + t\boldsymbol{Y}\right).$$

Remark. *Note that if we think of $\boldsymbol{X}$ and $\boldsymbol{Y}$ each as a vector of numbers, then this is the (matrix-valued) directional derivative of f at $\boldsymbol{X}$ in the direction of $\boldsymbol{Y}$.*

Consider $f(\boldsymbol{X}) = \boldsymbol{X}^{-1}$ for an invertible matrix $\boldsymbol{X} \in \mathbb{R}^{n\times n}$. Prove that

$$Df(\boldsymbol{X})[\boldsymbol{Y}] = -\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}.$$

Hint: You might need to use Exercise 1.

Solution.

We can write $\boldsymbol{Y} = \sum_{i,j}\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top$, where $\boldsymbol{e}_i$ denotes the vector with a 1 in the i-th coordinate and 0's elsewhere. Thus, we have that

$$Df\left(\boldsymbol{X}\right)\left[\boldsymbol{Y}\right] = \sum_{i,j} Df\left(\boldsymbol{X}\right)\left[\boldsymbol{Y}\left(i,j\right)\boldsymbol{e}_i\boldsymbol{e}_j^\top\right].$$

Therefore, if we prove that $Df(\boldsymbol{X})[\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top] = -\boldsymbol{X}^{-1}\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}$ then we can conclude that $Df(\boldsymbol{X})[\boldsymbol{Y}] = -\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}$.

We know that

$$Df(\boldsymbol{X})\left[\boldsymbol{Y}(i,j)\,\boldsymbol{e}_i\boldsymbol{e}_j^\top\right] = \lim_{t\to 0}\frac{\left(\boldsymbol{X} + t\,\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\right)^{-1} - \boldsymbol{X}^{-1}}{t}. \tag{2}$$

Furthermore, by applying Exercise 3 we get

$$\left(\boldsymbol{X} + t\,\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\right)^{-1} = \boldsymbol{X}^{-1} - \frac{\boldsymbol{X}^{-1}t\,\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}}{1 + t\,\boldsymbol{Y}(i,j)\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}\boldsymbol{e}_i}. \tag{3}$$

Combining Equations (2) and (3) implies that

$$Df(\boldsymbol{X})\left[\boldsymbol{Y}(i,j)\,\boldsymbol{e}_i\boldsymbol{e}_j^\top\right] = \lim_{t\to 0}\frac{-\boldsymbol{X}^{-1}\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}}{1 + t\,\boldsymbol{Y}(i,j)\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}\boldsymbol{e}_i} = -\boldsymbol{X}^{-1}\boldsymbol{Y}(i,j)\boldsymbol{e}_i\boldsymbol{e}_j^\top\boldsymbol{X}^{-1}.$$

This finishes the proof.

Exercise 3

1. Consider $\boldsymbol{A} \in S_{++}^n$ and matrix $\boldsymbol{\Delta} \in S_+^n$. Prove that $(\boldsymbol{A} + \boldsymbol{\Delta})^{-1} \preceq \boldsymbol{A}^{-1}$.

2. Let T be a convex set. We say that a function $f : T \to \mathbb{R}^{n\times n}$, is operator convex if for any two matrices $\boldsymbol{A}, \boldsymbol{B} \in T$ and any $\theta \in [0,1]$

$$f(\theta\boldsymbol{X} + (1-\theta)\,\boldsymbol{Y}) \preceq \theta f(\boldsymbol{X}) + (1-\theta)\,f(\boldsymbol{Y}).$$

   Prove that $f(\boldsymbol{X}) = \boldsymbol{X}^{-1}$ is operator convex over the set $T = S_{++}^n$.

   Hint: You could first show that operator convexity is implied by the second directional derivative $D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]$ being positive semi-definite for all $\boldsymbol{Y} \in S^n$ and $\boldsymbol{X} \in S_{++}^n$.

Solution.

1. From Exercise 1 in problem set 5, we know that if $\boldsymbol{A} \preceq \boldsymbol{B}$ for two matrices $\boldsymbol{A}, \boldsymbol{B} \in S_{++}^n$, then $\boldsymbol{B}^{-1} \preceq \boldsymbol{A}^{-1}$. By setting $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{\Delta}$, we can conclude that $(\boldsymbol{A} + \boldsymbol{\Delta})^{-1} \preceq \boldsymbol{A}^{-1}$.

   Note that by setting $\boldsymbol{\Delta} = \boldsymbol{B} - \boldsymbol{A}$, we can prove the statement of Exercise 1 from problem set 5. Thus, it would be interesting to provide an alternative proof which does not use this exercise. We provide such an alternative proof.

   We know that

$$(\boldsymbol{A} + \boldsymbol{\Delta})^{-1} = \boldsymbol{A}^{-1} + \int_{t=0}^{1}\frac{d}{dt}(\boldsymbol{A} + t\boldsymbol{\Delta})^{-1}\,dt.$$

   By applying Exercise 4, we get

$$(\boldsymbol{A} + \boldsymbol{\Delta})^{-1} = \boldsymbol{A}^{-1} + \int_{t=0}^{1} -(\boldsymbol{A} + t\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}(\boldsymbol{A} + t\boldsymbol{\Delta})^{-1}\,dt.$$

Consider an arbitrary vector $\boldsymbol{x}$, then we have that

$$\boldsymbol{x}^\top (\boldsymbol{A}+\boldsymbol{\Delta})^{-1} \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{A}^{-1}\boldsymbol{x} - \int_{t=0}^{1} \boldsymbol{x}^\top (\boldsymbol{A}+t\boldsymbol{\Delta})^{-1} \boldsymbol{\Delta} (\boldsymbol{A}+t\boldsymbol{\Delta})^{-1} \boldsymbol{x}\ dt.$$

We observe that $\boldsymbol{x}^\top (\boldsymbol{A}+t\boldsymbol{\Delta})^{-1} \boldsymbol{\Delta} (\boldsymbol{A}+t\boldsymbol{\Delta})^{-1} \boldsymbol{x} \geq 0$. Therefore, we can conclude that

$$\boldsymbol{x}^\top (\boldsymbol{A}+\boldsymbol{\Delta})^{-1} \boldsymbol{x} \leq \boldsymbol{x}^\top \boldsymbol{A}^{-1}\boldsymbol{x}$$

which implies that $(\boldsymbol{A}+\boldsymbol{\Delta})^{-1} \preceq \boldsymbol{A}^{-1}$.

2. First, we will show that $D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]$ being positive semi-definite for all $\boldsymbol{Y} \in S^n$ and $\boldsymbol{X} \in S^n_{++}$ implies operator-convexity. Define for $t \in [0,1]$ a function $h(t) = \boldsymbol{z}^\top f(\boldsymbol{X} + t(\boldsymbol{Y}-\boldsymbol{X}))\boldsymbol{z}$ for fixed $\boldsymbol{z} \in \mathbb{R}^n$ and $\boldsymbol{X}, \boldsymbol{Y} \in S^n_{++}$. This is a valid definition since $\boldsymbol{X} + t(\boldsymbol{Y}-\boldsymbol{X})$ is in S^n_{++} for $t \in [0,1]$. Observe that:

$$\frac{\partial}{\partial t} h(t) = \boldsymbol{z}^\top Df(\boldsymbol{X} + t(\boldsymbol{Y}-\boldsymbol{X}))[\boldsymbol{Y}-\boldsymbol{X}]\boldsymbol{z}$$
$$\frac{\partial^2}{\partial t^2} h(t) = \boldsymbol{z}^\top D^2 f(\boldsymbol{X} + t(\boldsymbol{Y}-\boldsymbol{X}))[\boldsymbol{Y}-\boldsymbol{X}, \boldsymbol{Y}-\boldsymbol{X}]\boldsymbol{z} \geq 0$$

The last inequality follows from $D^2 f(\boldsymbol{X} + t(\boldsymbol{Y}-\boldsymbol{X}))[\boldsymbol{Y}-\boldsymbol{X}, \boldsymbol{Y}-\boldsymbol{X}]$ being positive semi-definite by assumption. Therefore, we know that h is convex:

$$\boldsymbol{z}^\top [\theta f(\boldsymbol{X}) + (1-\theta) f(\boldsymbol{Y})]\, \boldsymbol{z} = \theta h(0) + (1-\theta) h(1) \geq h(1-\theta) = \boldsymbol{z}^\top f(\theta \boldsymbol{X} + (1-\theta)\boldsymbol{Y})\boldsymbol{z}$$

This inequality holds for any $\boldsymbol{z} \in \mathbb{R}^n$ which proves that f is operator convex.

What is left to show is that $D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]$ is indeed positive semi-definite. Using Exercise 2:

$$\begin{aligned}
D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}] &= \frac{\partial^2}{\partial t^2}\Big|_{t=0} f(\boldsymbol{X} + t\boldsymbol{Y}) \\
&= \frac{\partial}{\partial t}\Big|_{t=0} - (\boldsymbol{X} - t\boldsymbol{Y})^{-1} \boldsymbol{Y} (\boldsymbol{X} - t\boldsymbol{Y})^{-1} \\
&= -\frac{\partial (\boldsymbol{X} - t\boldsymbol{Y})^{-1}}{\partial t}\Big|_{t=0} \boldsymbol{Y}\boldsymbol{X}^{-1} - \boldsymbol{X}^{-1}\boldsymbol{Y} \frac{\partial (\boldsymbol{X} - t\boldsymbol{Y})^{-1}}{\partial t}\Big|_{t=0} \\
&= \boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1} + \boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1} \\
&= 2\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}
\end{aligned}$$

Remember that $\boldsymbol{X}$ being positive definite implies $\boldsymbol{X}^{-1}$ being positive definite. Thus, we have for any $\boldsymbol{x} \in \mathbb{R}^n$:

$$\boldsymbol{x}^\top D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]x = 2\boldsymbol{x}^\top \boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{x} = 2\boldsymbol{z}^\top \boldsymbol{X}^{-1}\boldsymbol{z} \geq 0$$

where $\boldsymbol{z} = \boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{x}$. Hence, $D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]$ is positive semi-definite which completes the proof.

Exercise 4: Max Flow in directed Graphs with Edge Capacities

Consider directed graph $G = (V, E, \boldsymbol{c})$ with arbitrary capacities $\boldsymbol{c} \geq \boldsymbol{0}$.

Let $\boldsymbol{B} \in \mathbb{R}^{E \times V}$ be the edge vertex incidence matrix of the graph, i.e. if $e \in E$ and $(u, v) = e$ then $\boldsymbol{B}(e, u) = 1$ and $\boldsymbol{B}(e, v) = -1$.

We let $\boldsymbol{\chi}_v \in \mathbb{R}^V$ denote the indicator of vertex v, i.e. $\boldsymbol{\chi}_v(v) = 1$ and $\boldsymbol{\chi}_v(u) = 0$ for $u \neq v$.

We let $s \in V$ denote the flow "source" and $t \in V$ the flow "sink".

The maximum flow problem is given by

$$\begin{aligned} \max_{\boldsymbol{f} \in \mathbb{R}^E, F \geq 0} \quad & F \\ \text{s.t. } \boldsymbol{B}\boldsymbol{f} = F(-\boldsymbol{\chi}_s + \boldsymbol{\chi}_t) & \\ \boldsymbol{0} \leq \boldsymbol{f} \leq \boldsymbol{c} & \end{aligned}$$

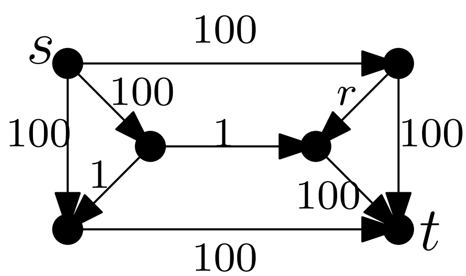
In the context of a given maximum flow problem, for a flow $\boldsymbol{f}$ satisfying $\boldsymbol{B}\boldsymbol{f} = F(-\boldsymbol{\chi}_s + \boldsymbol{\chi}_t)$, we define $\text{val}(\boldsymbol{f}) = F$.

Let $\boldsymbol{f}^*$ denote a feasible flow maximizing F, so that the maximum attainable flow value F is $\text{val}(\boldsymbol{f}^*)$.

Exercise 4.A: Convergence of Ford-Fulkerson

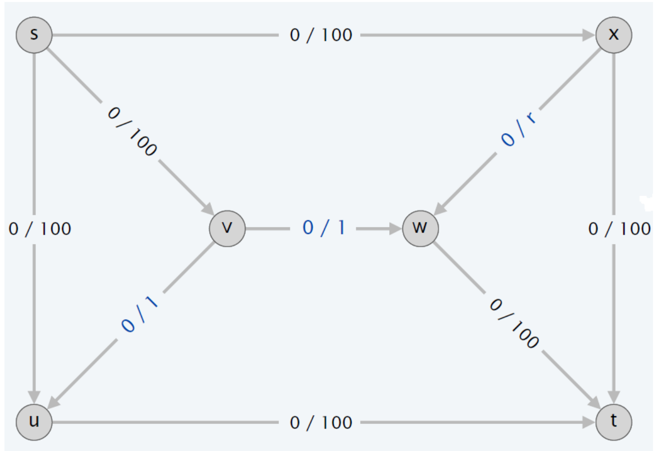
Show that the Ford-Fulkerson algorithm may not terminate; moreover, it may converge a value not equal to the value of the maximum flow.

Hint: You might use the graph below with the given capacities, where $r = \frac{\sqrt{5}-1}{2}$ (which implies that $r^2 = 1 - r$).

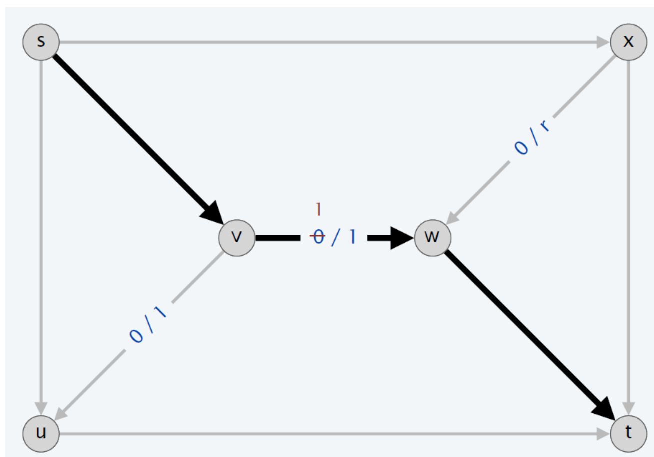


Solution. Consider the graph below with the given capacities.[1] The goal is to find a sequence of augmenting paths such that the algorithm converges to a value which is much smaller than the value of maximum flow.

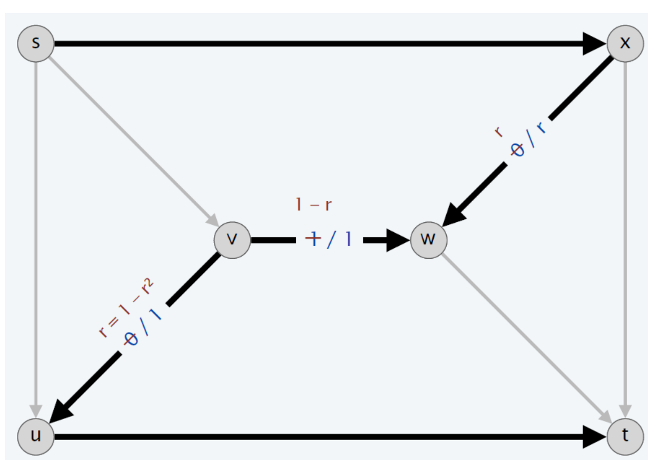
[1] All the figures used in this exercise are borrowed from the slides of the course Theory of Algorithms by Kevin Wayne, Princeton University.



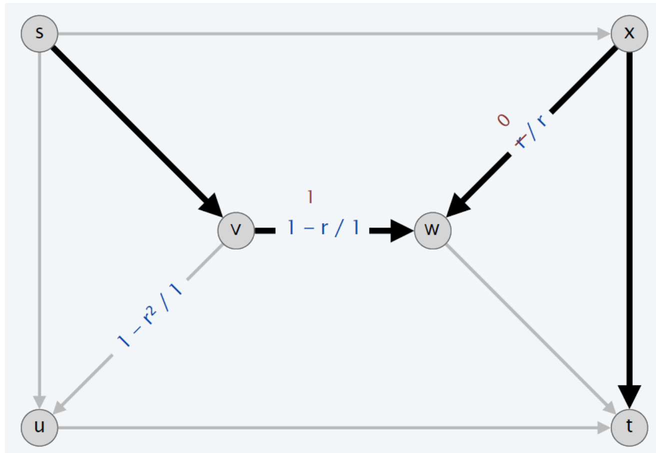
First, we augment Path 1: s, v, w, t.



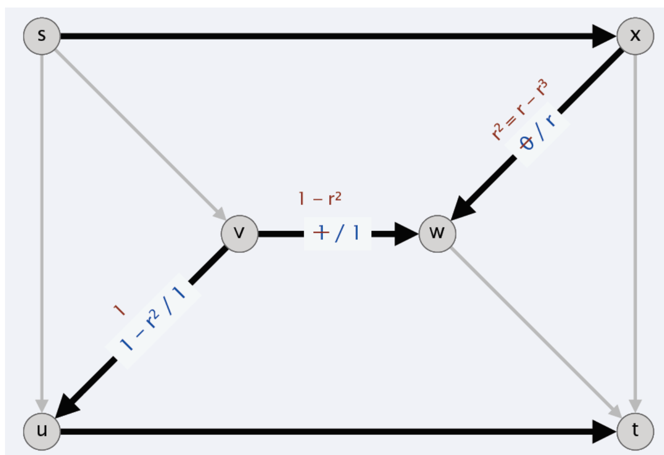
Then, we augment Path 2: s, x, w, v, u, t.



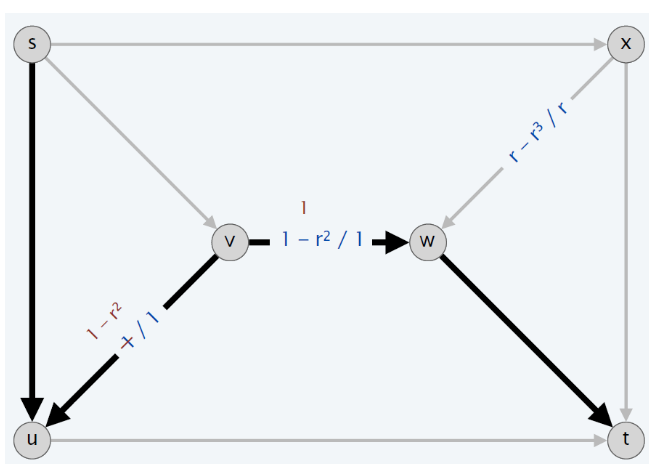
Then, we augment Path 3: s, v, w, x, t.



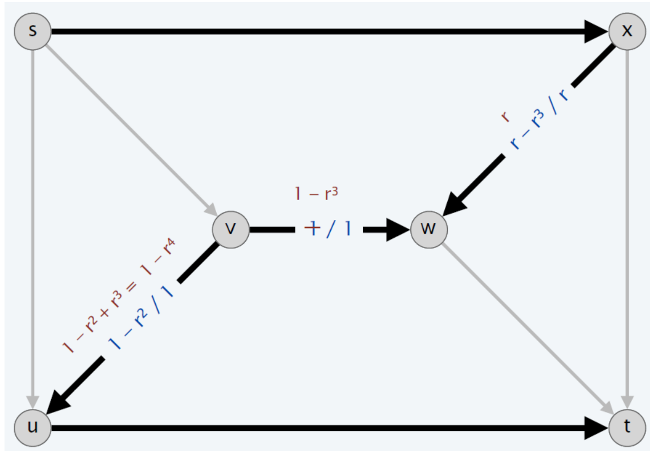
Then, we augment Path 4: s, x, w, v, u, t.



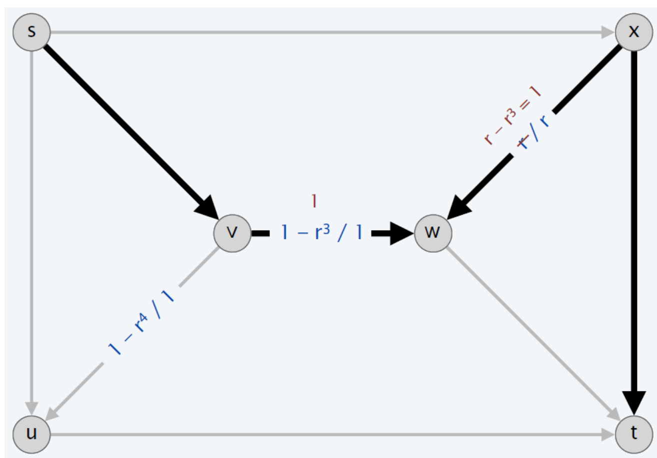
Then, we augment Path 5: s, u, v, w, t.



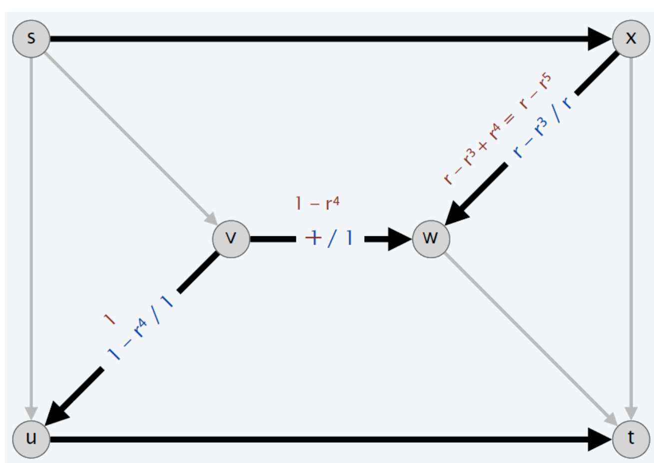
Then, we augment Path 6: s, x, w, v, u, t.



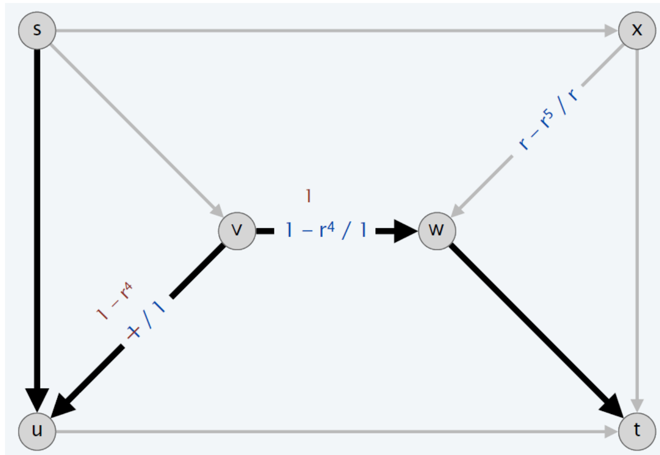
Then, we augment Path 7: s, v, w, x, t.



Then, we augment Path 8: s, x, w, v, u, t.



Then, we augment Path 9: s, u, v, w, t.



After augmenting Path 1 the flow is 1.

After augmenting Path 5 the flow is $1 + 2r + 2r^2$.

After augmenting Path 9 the flow is $1 + 2r + 2r^2 + 2r^3 + 2r^4$.

Using the given sequence of augmenting paths, after $(1+4k)$-th such path, the value of the flow is equal to

$$F = 1 + 2\sum_{i=1}^{k} r^i \leq 1 + 2\sum_{i=1}^{\infty} r^i = 3 + 2r \leq 5. \tag{4}$$

We observe that the value of maximum flow is equal to $200 + 1 = 201$. Therefore, we can conclude that the Ford-Fulkerson algorithm may not terminate; moreover, it may converge a value not equal to the value of the maximum flow.

Exercise 4.B: Iterative Refinement for Maximum Flow

Suppose we have an algorithm FlowRefine, which given a maximum flow instance $G = (V, E, \boldsymbol{c})$ with source $s \in V$ and sink $t \in V$ returns a feasible s-t flow $\tilde{\boldsymbol{f}}$, i.e. $\boldsymbol{B}\tilde{\boldsymbol{f}} = F(-\boldsymbol{\chi}_s + \boldsymbol{\chi}_t)$ for some F, and $\boldsymbol{0} \leq \tilde{\boldsymbol{f}} \leq \boldsymbol{c}$, and $\tilde{\boldsymbol{f}}$ is guaranteed to route at least half the maximum flow, i.e. $F = \text{val}(\tilde{\boldsymbol{f}}) \geq 0.5\,\text{val}(\boldsymbol{f}^*)$.

Suppose that the running time of FlowRefine is $O(|E|^c)$ for some constant $c \geq 1$.

Explain how we can use FlowRefine to find a flow $\hat{\boldsymbol{f}}$ that routes at least $(1-\epsilon)\,\text{val}(\boldsymbol{f}^*)$ in time $O(|E|^c \log(1/\epsilon))$.

Solution. Assume that we are given a maximum flow instance $G = (V, E, \boldsymbol{c})$ with source $s \in V$ and sink $t \in V$. Consider the flow $\boldsymbol{f} = \boldsymbol{0}$. We run the algorithm FlowRefine on graph G. We know that, it will return a feasible s-t flow $\boldsymbol{f}_1$ such that $\text{val}(\boldsymbol{f}_1) \geq 0.5\,\text{val}(\boldsymbol{f}^*_G)$, where $\boldsymbol{f}^*_G$ is a maximum flow in G. Let $G_{\boldsymbol{f}_1}$ be the residual graph of $\boldsymbol{f}_1$ and set $\boldsymbol{f} = \boldsymbol{f} + \boldsymbol{f}_1$. Let's run FlowRefine this time on graph $G_{\boldsymbol{f}_1}$. It should return a flow $\boldsymbol{f}_2$. Now, we set $\boldsymbol{f} = \boldsymbol{f} + \boldsymbol{f}_2$. Thus, we get

$$\begin{aligned}
\mathrm{val}(\boldsymbol{f}) &= \mathrm{val}(\boldsymbol{f}_1) + \mathrm{val}(\boldsymbol{f}_2) \\
&\geq \mathrm{val}(\boldsymbol{f}_1) + \frac{1}{2}\mathrm{val}\left(\boldsymbol{f}^*_{G_{f_1}}\right) \\
&= \mathrm{val}(\boldsymbol{f}_1) + \frac{1}{2}\mathrm{val}(\boldsymbol{f}^*_G) - \frac{1}{2}\mathrm{val}(\boldsymbol{f}_1) \\
&= \frac{1}{2}\mathrm{val}(\boldsymbol{f}_1) + \frac{1}{2}\mathrm{val}(\boldsymbol{f}^*_G) \\
&\geq \frac{1}{4}\mathrm{val}(\boldsymbol{f}^*_G) + \frac{1}{2}\mathrm{val}(\boldsymbol{f}^*_G) \\
&\geq \frac{3}{4}\mathrm{val}(\boldsymbol{f}^*_G)
\end{aligned}$$

where we used $\mathrm{val}(\boldsymbol{f}^*_{G_{f_1}}) + \mathrm{val}(\boldsymbol{f}_1) = \mathrm{val}(\boldsymbol{f}^*_G)$ based on Lemmas 2.6 and 2.7 from Lecture 10.

By repeating the same argument, we can conclude that after $k = \log(1/\epsilon)$ iterations

$$\mathrm{val}(\boldsymbol{f}) \geq \left(1 - \frac{1}{2^k}\right)\mathrm{val}(\boldsymbol{f}^*_G) = (1-\epsilon)\,\mathrm{val}(\boldsymbol{f}^*_G)$$

We know that the running time of FlowRefine is $O(|E|^c)$ for some constant $c \geq 1$ and one can generate a residual graph in $O(|E|)$. Therefore, the above algorithm runs in time $O(|E|^c \log(1/\epsilon))$.

Exercise 5

Maximum flow is often introduced as a s-t problem in the following manner: Given a directed and capacitated graph $G = (V, E, \boldsymbol{c})$ with n vertices and m edges (assume integer capacities between 1 and n^{10}) and edge-vertex incidence matrix $\boldsymbol{B}$, a source $s \in V$ and a sink $t \in V$, computes the flow $\boldsymbol{f}$ that respects the capacities $\boldsymbol{0} \leq \boldsymbol{f} \leq \boldsymbol{c}$, satisfies flow conservation $(\boldsymbol{Bf})(v) =$ for $v \in V \setminus \{s, t\}$ and maximizes F where $(\boldsymbol{Bf})(s) = -F$ and $(\boldsymbol{Bf})(t) = F$.

1. Show that such an algorithm can be used to solve the maximum flow problem as introduced in the lecture, i.e. given a integral demand vector $\boldsymbol{d} \perp \boldsymbol{1}$ with entries between $-n^{10}$ and n^{10}, compute a flow that satisfies
$$\min_{\boldsymbol{0} \leq \boldsymbol{f}, \boldsymbol{Bf} = \boldsymbol{d}} \left\| \mathrm{diag}(\boldsymbol{c})^{-1} \boldsymbol{f} \right\|_\infty . \tag{5}$$
2. Show that the converse is also true, i.e. that a solution to (5) can be used to solve the s-t maximum flow problem.

Hint: There is always an integral maximum flow $\boldsymbol{f}$ *in this setting.*

Solution.

1. To compute such a flow using s-t maximum flow, we add extra vertices s and t, and connect every vertex v with negative demand $\boldsymbol{d}(v)$ to s with an edge (s, v) of capacity $-\boldsymbol{d}(v)$. We do the converse for every vertex u with positive demand, i.e. we add an edge (u, t) of capacity

$\boldsymbol{d}(u)$. Then, we finally aim to find the minimum necessary congestion to route the demands. To do so, we scale all the congestions in the input graph with some factor α and then round them to the nearest integer. Binary searching for the lowest factor α that admits a flow yields the congestion minimizing flow that routes the demands. Notice that the search can be terminated once α is sufficiently (polynomially in n) accurate, because the optimal flow is integral.

2. To compute such a flow using s-t maximum flow, we add extra vertices s and t, and connect every vertex v with negative demand $\boldsymbol{d}(v)$ to s with an edge (s, v) of capacity $-\boldsymbol{d}(v)$. We do the converse for every vertex u with positive demand, i.e. we add an edge (u, t) of capacity $\boldsymbol{d}(u)$. Then, we finally aim to find the minimum necessary congestion to route the demands. To do so, we scale all the congestions in the input graph with some factor α and then round them to the nearest integer. Binary searching for the lowest factor α that admits a flow routing all the demands yields the congestion minimizing flow that routes the demands. Notice that the search can be terminated once α is sufficiently (polynomially in n) accurate, because the optimal flow is integral.

3. To compute compute the maximum flow, we again use the oracle with demands -1 and 1 at s and t respectively. Then, we can scale the flow with $1/\left\|\mathrm{diag}(\boldsymbol{c})^{-1}\boldsymbol{f}\right\|_\infty$.