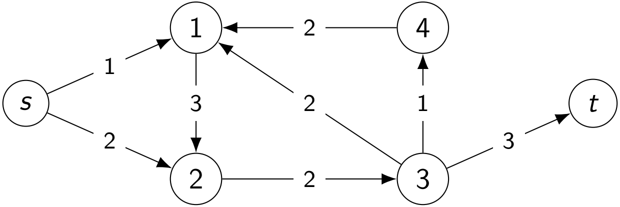
Network Flows

Math 482, Lecture 23

Misha Lavrov

March 30, 2020

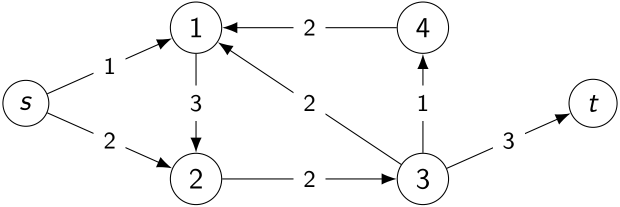
Definition of a network



- A set N of *nodes*. Here, $N = \{s, 1, 2, 3, 4, t\}$.

  Node s is the *source* and node t is the *sink*.

Definition of a network



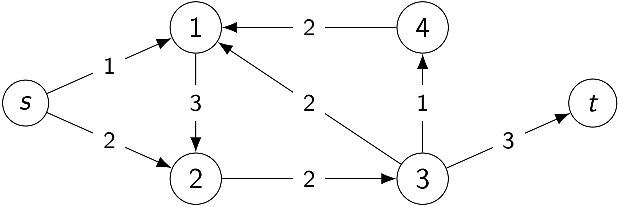
- A set N of *nodes*. Here, $N = \{s, 1, 2, 3, 4, t\}$.

  Node s is the *source* and node t is the *sink*.

- A set A of *arcs*: pairs of nodes. In this example,

  $A = \{(s, 1), (s, 2), (1, 2), (2, 3), (3, 1), (3, 4), (3, t), (4, 1)\}$.

Definition of a network



- A set N of *nodes*. Here, $N = \{s, 1, 2, 3, 4, t\}$.

  Node s is the *source* and node t is the *sink*.

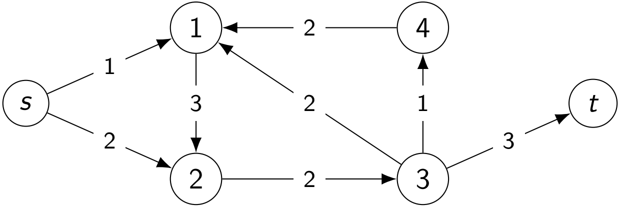
- A set A of *arcs*: pairs of nodes. In this example,

  $A = \{(s, 1), (s, 2), (1, 2), (2, 3), (3, 1), (3, 4), (3, t), (4, 1)\}$.

- A nonnegative real capacity c_{ij} on each arc (i, j).

  Here, $c_{s1} = 1$, $c_{12} = 3$, and so on.

Definition of a network



- A set N of *nodes*. Here, $N = \{s, 1, 2, 3, 4, t\}$.

  Node s is the *source* and node t is the *sink*.

- A set A of *arcs*: pairs of nodes. In this example,

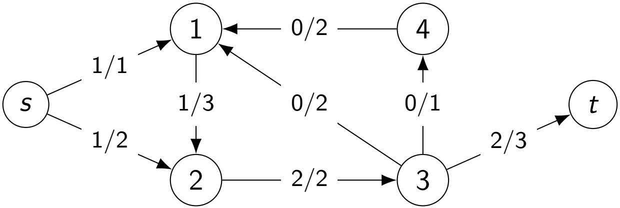
  $A = \{(s, 1), (s, 2), (1, 2), (2, 3), (3, 1), (3, 4), (3, t), (4, 1)\}$.

- A nonnegative real capacity c_{ij} on each arc (i, j).

  Here, $c_{s1} = 1$, $c_{12} = 3$, and so on.

Definition of a network flow, I

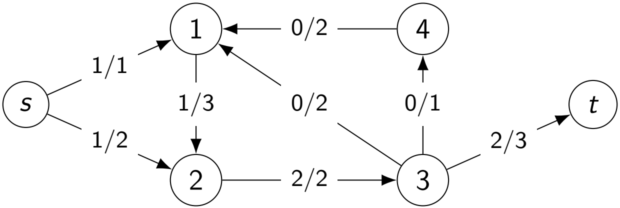
A flow **x** assigns a number x_{ij} to each arc $(i, j) \in A$.



- We write "p/q" on an arc (i, j) with flow $x_{ij} = p$ and capacity $c_{ij} = q$.

Definition of a network flow, I

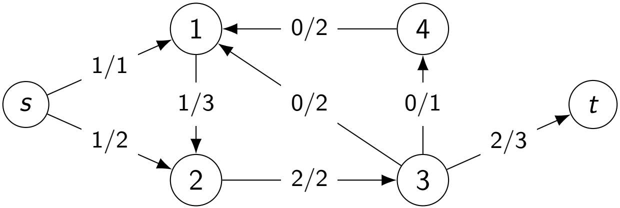
A flow $\mathbf{x}$ assigns a number x_{ij} to each arc $(i, j) \in A$.



- We write "p/q" on an arc (i, j) with flow $x_{ij} = p$ and capacity $c_{ij} = q$.
- A flow represents stuff moving from s to t; x_{ij} is the amount of stuff moving along arc (i, j).

Definition of a network flow, I

A flow $\mathbf{x}$ assigns a number x_{ij} to each arc $(i,j) \in A$.



- We write "p/q" on an arc (i,j) with flow $x_{ij} = p$ and capacity $c_{ij} = q$.
- A flow represents stuff moving from s to t; x_{ij} is the amount of stuff moving along arc (i,j).
- For this to make sense, we want to add some constraints on $\mathbf{x}$ for it to be a feasible flow.

Definition of a network flow, II

Constraints on a feasible flow:

- Capacity constraints: for every arc $(i, j) \in A$, $x_{ij} \leq c_{ij}$.

Definition of a network flow, II

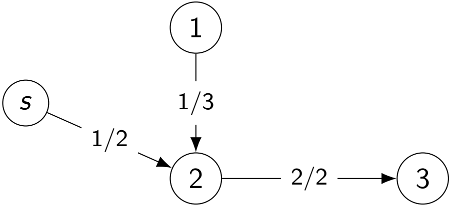
Constraints on a feasible flow:

- Capacity constraints: for every arc $(i, j) \in A$, $x_{ij} \leq c_{ij}$.
- Nonnegativity constraints: $\mathbf{x} \geq \mathbf{0}$.

Definition of a network flow, II

Constraints on a feasible flow:

- Capacity constraints: for every arc $(i, j) \in A$, $x_{ij} \le c_{ij}$.
- Nonnegativity constraints: $\mathbf{x} \ge \mathbf{0}$.
- Flow conservation: at every node $k \in N$ **except for** s **and** t, the total flow going in is equal to the total flow going out.



At node $k = 2$, we must have $x_{s2} + x_{12} = x_{23}$. Here, $1 + 1 = 2$.

More on flow conservation

The *excess* at a node k is the difference between the total flow into k and the total flow out of k:

$$\Delta_k(\mathbf{x}) := \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj}.$$

More on flow conservation

The *excess* at a node k is the difference between the total flow into k and the total flow out of k:

$$\Delta_k(\mathbf{x}) := \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj}.$$

Flow conservation: at every node $k \neq s, t$, $\Delta_k(\mathbf{x}) = 0$.

More on flow conservation

The *excess* at a node k is the difference between the total flow into k and the total flow out of k:

$$\Delta_k(\mathbf{x}) := \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj}.$$

Flow conservation: at every node $k \neq s, t$, $\Delta_k(\mathbf{x}) = 0$.

At the sink t, the excess $\Delta_t(\mathbf{x})$ represents the amount of stuff we've successfully brought to t from s. This is called the *value* of $\mathbf{x}$.

More on flow conservation

The *excess* at a node k is the difference between the total flow into k and the total flow out of k:

$$\Delta_k(\mathbf{x}) := \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj}.$$

Flow conservation: at every node $k \neq s, t$, $\Delta_k(\mathbf{x}) = 0$.

At the sink t, the excess $\Delta_t(\mathbf{x})$ represents the amount of stuff we've successfully brought to t from s. This is called the *value* of $\mathbf{x}$.

We can prove that $\Delta_s(\mathbf{x}) = -\Delta_t(\mathbf{x})$: the amount of gain at t is equal to the amount of loss at s. (This should follow from flow conservation.)

The maximum flow LP

The maximum flow problem to find the feasible flow in a network with the maximum value can be written as a linear program:

$$
\begin{aligned}
\underset{\mathbf{x}\in\mathbb{R}^{|A|}}{\text{maximize}} \quad & \sum_{i:(i,t)\in A} x_{it} - \sum_{j:(t,j)\in A} x_{tj} \\
\text{subject to} \quad & \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj} = 0 \quad (k \in N, k \neq s, t) \\
& x_{ij} \leq c_{ij} \qquad (i,j) \in A \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}
$$

The maximum flow LP

The maximum flow problem to find the feasible flow in a network with the maximum value can be written as a linear program:

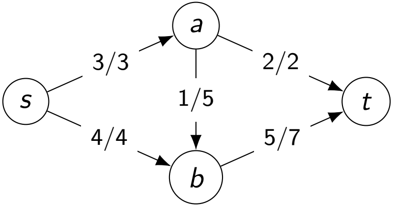
$$\begin{aligned}
\underset{\mathbf{x}\in\mathbb{R}^{|A|}}{\text{maximize}} \quad & \sum_{i:(i,t)\in A} x_{it} - \sum_{j:(t,j)\in A} x_{tj} \\
\text{subject to} \quad & \sum_{i:(i,k)\in A} x_{ik} - \sum_{j:(k,j)\in A} x_{kj} = 0 \quad (k \in N, k \neq s, t) \\
& x_{ij} \leq c_{ij} \qquad (i,j) \in A \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}$$

We can assume there are no arcs into s or out of t. In that case,

$$\text{value of } \mathbf{x} = \sum_{i:(i,t)\in A} x_{it} = \sum_{j:(s,j)\in A} x_{sj}.$$

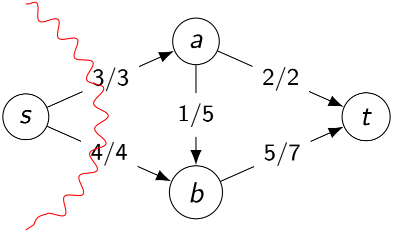
How can we tell if a flow is optimal?

The flow in the example below has a value of 7. Can we do better?



How can we tell if a flow is optimal?

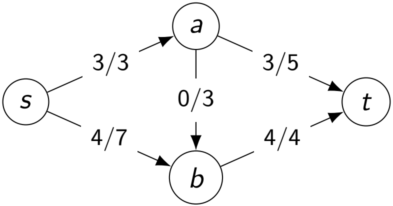
The flow in the example below has a value of 7. Can we do better?



No; the arcs out of s are all at their maximum capacity. We can't send more than 7 flow out of s.

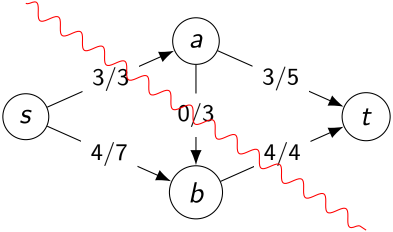
How can we tell if a flow is optimal?

The flow in the example below has a value of 7. Can we do better?



How can we tell if a flow is optimal?

The flow in the example below has a value of 7. Can we do better?



No: the arcs from $\{s, b\}$ to $\{a, t\}$ are all at their maximum capacity, and the arcs from $\{a, t\}$ to $\{s, b\}$ are all at capacity 0. We can't send more than 7 flow from $\{s, b\}$ to $\{a, t\}$.

Cuts

Definition

An *cut* in a network is a partition of the node set N into two sets S and T, such that $s \in S$ and $t \in T$.

Cuts

Definition

An *cut* in a network is a partition of the node set N into two sets S and T, such that $s \in S$ and $t \in T$.

The *capacity* of a cut (S, T) is the sum $\sum_{i \in S} \sum_{j \in T} c_{ij}$.

(If $(i, j) \notin A$, we say that $c_{ij} = 0$.)

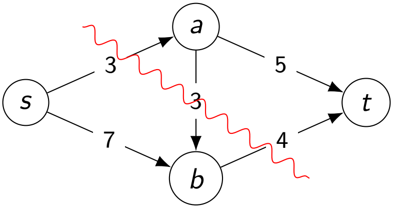
Cuts

Definition

An *cut* in a network is a partition of the node set N into two sets S and T, such that $s \in S$ and $t \in T$.

The *capacity* of a cut (S, T) is the sum $\sum_{i \in S} \sum_{j \in T} c_{ij}$.

(If $(i, j) \notin A$, we say that $c_{ij} = 0$.)



Here, $S = \{s, b\}$, $T = \{a, t\}$, and the capacity is $c_{sa} + c_{bt} = 7$.

Cuts are upper bounds on flows

Theorem

If a feasible flow $\mathbf{x}$ *has value* $v(\mathbf{x})$*, and a cut* (S, T) *has capacity* $c(S, T)$*, then*

$$v(\mathbf{x}) \leq c(S, T).$$

Cuts are upper bounds on flows

Theorem

If a feasible flow $\mathbf{x}$ *has value* $v(\mathbf{x})$*, and a cut* (S, T) *has capacity* $c(S, T)$*, then*

$$v(\mathbf{x}) \leq c(S, T).$$

Proof idea: consider the sum

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right)$$

By computing this sum in two ways, we show that it is equal to $v(\mathbf{x})$, and also that it is at most $c(S, T)$.

Step 1

In the sum

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right)$$

the difference (in orange) is the net flow out of k. When $k \neq s$, it is 0. When $k = s$, it is the value of the flow.

Therefore

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right) = \sum_{j:(s,j) \in A} x_{sj} - \sum_{i:(i,s) \in A} x_{is} = v(\mathbf{x}).$$

Step 2

How many times, and with what sign, does x_{ij} appear in the sum

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right) ?$$

Step 2

How many times, and with what sign, does x_{ij} appear in the sum

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right) ?$$

- Once, with $+$ sign, if $i \in S$.
- Once, with $-$ sign, if $j \in S$.

Step 2

How many times, and with what sign, does x_{ij} appear in the sum

$$\sum_{k\in S}\left(\sum_{j:(k,j)\in A} x_{kj} - \sum_{i:(i,k)\in A} x_{ik}\right)?$$

- Once, with $+$ sign, if $i \in S$.
- Once, with $-$ sign, if $j \in S$.

Therefore

$$\sum_{k\in S}\left(\sum_{j:(k,j)\in A} x_{kj} - \sum_{i:(i,k)\in A} x_{ik}\right) = \sum_{i\in S}\sum_{j\in T} x_{ij} - \sum_{i\in T}\sum_{j\in S} x_{ij}.$$

Step 2, continued

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right) = \sum_{i \in S} \sum_{j \in T} x_{ij} - \sum_{i \in T} \sum_{j \in S} x_{ij}.$$

Step 2, continued

$$\sum_{k \in S} \left(\sum_{j:(k,j) \in A} x_{kj} - \sum_{i:(i,k) \in A} x_{ik} \right) = \sum_{i \in S} \sum_{j \in T} x_{ij} - \sum_{i \in T} \sum_{j \in S} x_{ij}.$$

For the sum in red, use $x_{ij} \leq c_{ij}$:

$$\sum_{i \in S} \sum_{j \in T} x_{ij} \leq \sum_{i \in S} \sum_{j \in T} c_{ij} = c(S, T).$$

Step 2, continued

$$\sum_{k\in S}\left(\sum_{j:(k,j)\in A} x_{kj} - \sum_{i:(i,k)\in A} x_{ik}\right) = \sum_{i\in S}\sum_{j\in T} x_{ij} - \sum_{i\in T}\sum_{j\in S} x_{ij}.$$

For the sum in red, use $x_{ij} \leq c_{ij}$:

$$\sum_{i\in S}\sum_{j\in T} x_{ij} \leq \sum_{i\in S}\sum_{j\in T} c_{ij} = c(S,T).$$

The sum in blue is ≥ 0, so for an upper bound, we can ignore it.

Step 2, continued

$$\sum_{k\in S}\left(\sum_{j:(k,j)\in A} x_{kj} - \sum_{i:(i,k)\in A} x_{ik}\right) = \sum_{i\in S}\sum_{j\in T} x_{ij} - \sum_{i\in T}\sum_{j\in S} x_{ij}.$$

For the sum in red, use $x_{ij} \leq c_{ij}$:

$$\sum_{i\in S}\sum_{j\in T} x_{ij} \leq \sum_{i\in S}\sum_{j\in T} c_{ij} = c(S,T).$$

The sum in blue is ≥ 0, so for an upper bound, we can ignore it.

We conclude that

$$v(\mathbf{x}) = \sum_{k\in S}\left(\sum_{j:(k,j)\in A} x_{kj} - \sum_{i:(i,k)\in A} x_{ik}\right) \leq c(S,T).$$