Operations Research (Master's Degree Course)

7.3 Problems on Graphs: Flows in Networks

Silvano Martello

DEI "Guglielmo Marconi", Università di Bologna, Italy



This work by is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

Based on a work at http://www.editrice-esculapio.com

Flows in Networks: The Maximum Flow Problem

- Network = Directed graph G = (V, A) having capacities associated with the arcs:
 - $q_{ij} =$ capacity of (v_i, v_j) (integer).
- ullet Arcs are seen as "pipes" in which some material can "flow": $q_{ij}=\mathsf{maximum}$ flow in (v_i,v_j) .
- Problem: given $s, t \in V$ (source and sink), send the maximum flow from s to t.
- Example: road map, with q_{ij} expressed in number of vehicles/hour; find the maximum quantity of traffic that is possible between two locations (and detect bottleneck roads).
- ullet Given G=(V,A) with $capacities\ q_{ij}$, a set of values

$$\xi_{ij}$$
 (= flow in (v_i, v_j) , $i, j = 1, \ldots, n$)

is called a **feasible flow of value** z **from** s **to** t if

$$0 \le \xi_{ij} \le q_{ij} \ \forall \ (v_i, v_j) \in A; \tag{\alpha}$$

$$\sum_{v_j \in \Gamma^+(v_i)} \xi_{ij} - \sum_{v_k \in \Gamma^-(v_i)} \xi_{ki} = \begin{cases} +z & \text{if } v_i = s \\ -z & \text{if } v_i = t \\ 0 & \text{otherwise} \end{cases} \quad \forall v_i \in V \quad (\beta)$$

- ullet An s-t cut is a partition of V in (V_1,V_2) : $s\in V_1$, $t\in V_2$.
- ullet The value of an s-t cut is $\sum_{(v_i,v_j)\ :\ v_i\in V_1,v_j\in V_2} q_{ij}.$

Max-flow min-cut theorem (Ford-Fulkerson, 1956)

The value z of the maximum flow from s to t is equal to the minimum value of an s-t cut.

Proof z cannot obviously be greater than the value of any s-t cut. We will show, through a **constructive proof** that a flow with such value exists.

ullet Given a feasible flow ξ_{ij} , let us execute the following procedure:

$$V_1:=\{s\};$$
 while $\exists~v_i\in V_1$ e $v_j
ot\in V_1:\xi_{ij}< q_{ij}$ or $\xi_{ji}>0$ do $V_1:=V_1\bigcup\{v_j\}.$

• Case 1: $t \in V_1$: $\Rightarrow \exists$ sequence of arcs connecting s to t

$$(S \longrightarrow V) \longrightarrow V) \longleftarrow V \longrightarrow V$$

such that: $\xi_{ij} < q_{ij} \; \forall \;$ forward arc (v_i, v_j) and $\xi_{kl} > 0 \; \forall \;$ backward arc (v_k, v_l) .

• Such a sequence is called an **augmenting chain**. Indeed:

$$\begin{array}{lll} \delta_1 & = & \min\{(q_{ij} - \xi_{ij}): & (v_i, v_j) \text{ is a forward arc}\} \\ \delta_2 & = & \min\{\xi_{kl}: & (v_k, v_l) \text{ is a backward arc}\} \end{array} \right\} \delta = \min(\delta_1, \delta_2). \blacksquare$$

adding (resp. subtracting) δ flow units to each forward (resp. backward) arc, new flow that

- satisfies $(\alpha) \leftarrow \text{definition of } \delta$
- has a value increased by δ units.
- ullet New flow o new augmenting chain ... until:

• At each iteration of Case 1 the flow increases by at least one unit \Rightarrow convergence. \Box

Number of iterations bounded by any upper bound on z, e.g., $U=\sum_{v_j\in\Gamma^+(s)}q_{sj}$

Time: $O(T \cdot U)$ (T = time needed by an iteration).

Algorithm outline:

- start with a feasible flow (e.g., $\xi_{ij}=0 \ \forall \ i,j$), and increase it through augmenting chains; when no augmenting chain exists, we have a maximum flow.
- For finding the augmenting chains we attach labels to the vertices. A vertex can be
 - unlabeled:
 - · labeled ($\Leftrightarrow \in V_1$) and unscanned;
 - · labeled and scanned (\Leftrightarrow used for trying to increase V_1).
- the label of vertex v_i has the form $\left\{ egin{array}{ll} [+v_k,\delta] &\Leftrightarrow &\xi_{ki} \mbox{ can be increased;} \\ [-v_k,\delta] &\Leftrightarrow &\xi_{ik} \mbox{ can be decreased;} \\ \end{array} \right.$ with $\delta=$ maximum additional flow that can be sent from s to v_i .

Ford-Fulkerson Algorithm

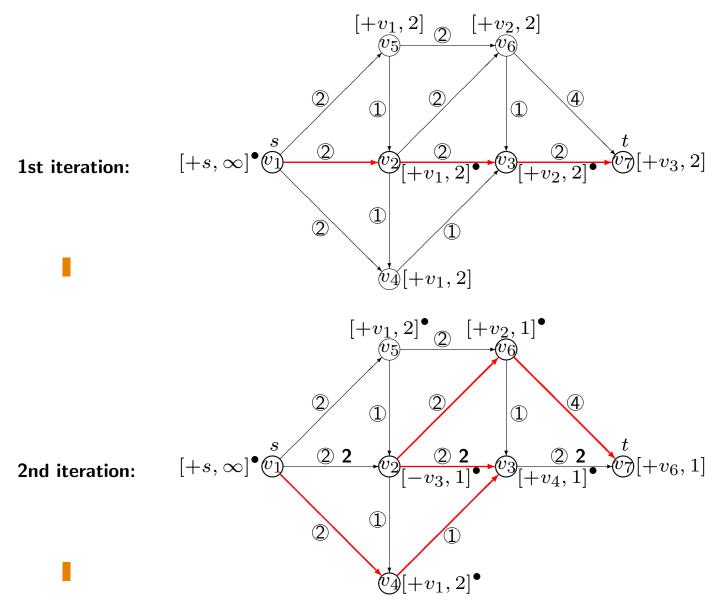
```
procedure MAX_FLOW:
begin
   for i := 1 to n do for j := 1 to n do \xi_{ij} := 0;
   opt := false;
   while opt = false do
      begin
          label s by [+s, +\infty];
          repeat
              let v_i be a labeled ([\pm v_k, \delta(v_i)]) unscanned vertex;
              for each v_i \in \Gamma^+(v_i) : v_i is unlabeled and \xi_{ij} < q_{ij} do
                 label v_i by [+v_i, \min(\delta(v_i), q_{ij} - \xi_{ij})];
             for each v_i \in \Gamma^-(v_i) : v_i is unlabeled and \xi_{ii} > 0 do
                 label v_i by [-v_i, \min(\delta(v_i), \xi_{ii})];
              mark v_i as scanned
          until t is labeled or no new vertex can be labeled:
          if t is unlabeled then opt := true
          else
```

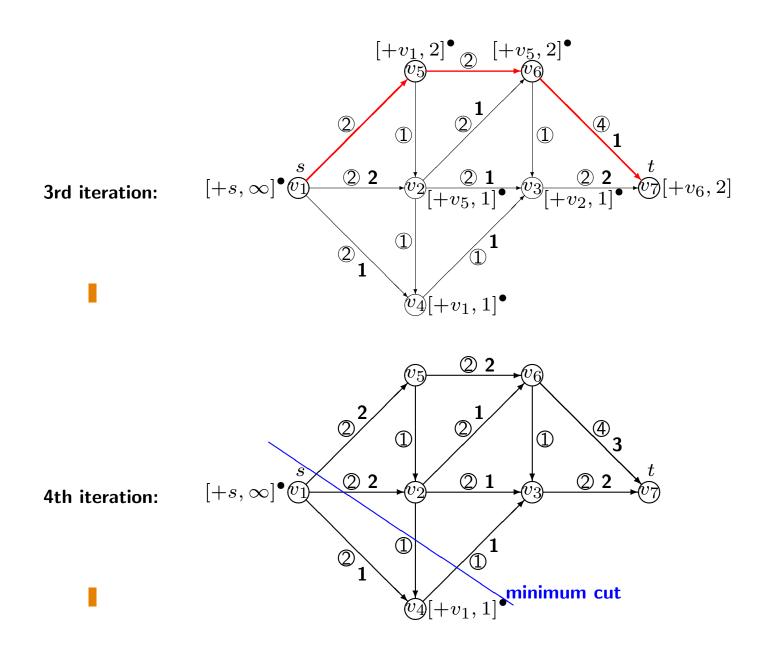
Ford-Fulkerson Algorithm (cont'd)

if t is unlabeled then opt := trueelse begin $\delta^* := \delta(t); \ x := t;$ repeat if the label of x is $[+y,\delta(x)]$ then $\xi_{yx}:=\xi_{yx}+\delta^*$ **else** (i.e., the label is $[-y, \delta(x)]$) $\xi_{xy} := \xi_{xy} - \delta^*$; x := yuntil x = s: cancel all labels end end: **comment**: the minimum cut is given by $V_1 = \{v_i : v_i \text{ is labeled}\}, \quad V_2 = V \setminus V_1$ end. In the **course web page**: **applet** for executing the Ford-Fulkerson algorithm. Remind: Number of iterations bounded by any upper bound on z, e.g., $U = \sum_{i=1}^n q_{sj} \Rightarrow$ $v_j \in \Gamma^+(s)$ time: $O(T \cdot U)$ (T= time needed by an iteration), pseudo-polynomial With this implementation, one iteration takes $O(n^2)$ time \Rightarrow overall time $O(n^2U)$.

Example: $s=v_1$, $t=v_7$. Capacities within circles. Start with nil flow.

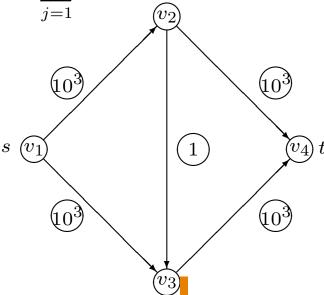
Vertices considered by increasing index. Symbol '•' marks the scanned vertices:





Complexity of maximum flow

- Ford-Fulkerson algorithm: $O(n^2)$ time per augmenting chain; at most z augmenting chains;
- upper bound on z: $\sum_{i=0}^{n} q_{sj} \Longrightarrow Complexity <math>O(n^2z)$, pseudo-polynomial. **Example:**



Possible sequence: 1st augmenting chain: v_1, v_2, v_3, v_4 ; $\delta = 1$;

2nd augmenting chain: v_1, v_3, v_2, v_4 ; $\delta = 1$;

3rd augmenting chain: v_1, v_2, v_3, v_4 ; $\delta = 1$;

... (2000 iterations)

- Edmonds and Karp (1972): by selecting, at each iteration, the **shortest augmenting chain**, at most n^3 iterations; $O(n^5)$ algorithm;
- Karzanov (1974): $O(n^3)$ algorithm.

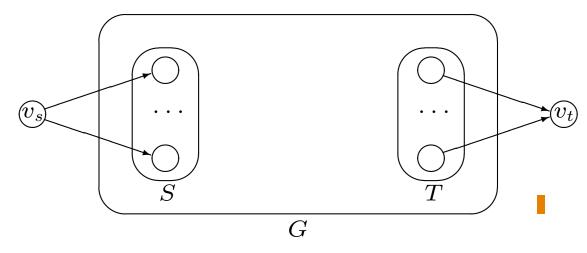
Other max-flow problems

Multiple sources and sinks:

- **Problem:** Given G = (V, A), with capacities q_{ij} , and S, $T \subset V$ $(S \cap T = \emptyset)$, send the maximum total flow from all sources of S to all sinks of T.
- Solution: define a new graph G' = (V', A'): with

$$V' := V \cup \{v_s, v_t\};$$

$$A' := A \cup \{(v_s, v_i) : v_i \in S\} \cup \{(v_j, v_t) : v_j \in T\}$$



and send the maximum flow from v_s to v_t .

• If there is no additional constraint, **capacities** $q_{si}=q_{jt}=+\infty \ \forall \ i,j,$ otherwise: $q_{si}=$ limit on the supply at source $v_i,\ q_{jt}=$ limit on the demand at sink $v_j.$

Other max-flow problems (cont'd)

Arc and vertex capacities:

• **Problem:**given G = (V, A), with $q_{ij} =$ capacity of arc (v_i, v_j) and $p_i =$ capacity of vertex v_i , and $s, t \in V$, send the maximum flow from s to t by satisfying the additional constraint

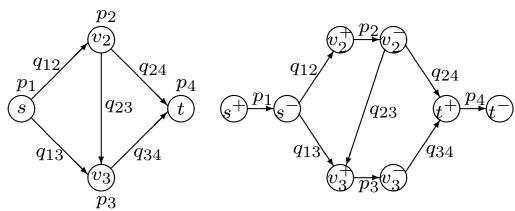
$$\sum_{v_k \in \Gamma^-(v_i)} \xi_{ki} \le p_i \ \forall \ v_i \in V \tag{\gamma}.$$

• Solution: new graph $G' = (V', A' \cup A'')$ with

$$V' := \{v_i^+ : v_i \in V\} \cup \{v_i^- : v_i \in V\};$$

$$A' = \{(v_i^-, v_j^+) : (v_i, v_j) \in A\}$$
 with capacity of $(v_i^-, v_j^+) = q_{ij}$;

$$A'' = \{(v_i^+, v_i^-) : v_i \in V\}$$
 with capacity of $(v_i^+, v_i^-) = p_i$



and send the maximum flow from s^+ to t^- .

• The total flow entering vertex v_i^+ must travel along arc $(v_i^+, v_i^-) \Rightarrow$ it must satisfy (γ) .

Flows in Networks: The Minimum Cost Flow Problem

- Max-Flow and Min-Cost Flow are the two most relevant flow problems.
- Network G = (V, A) having **two** positive integer values associated with each arc (v_i, v_j) :
 - capacity $q_{ij} = \text{maximum flow along } (v_i, v_j);$
 - cost c_{ij} = cost per unit of flow along (v_i, v_j) .
- ullet Problem: given $s,t\in V$, send a flow of prefixed value u from s to t at minimum cost.
- Algorithm: Phase 1. find a feasible flow of value u;

Phase 2. iteratively modify the flow preserving its value and decreasing its cost.

- Phase 1:
 - use the Ford-Fulkerson algorithm halting it when the current flow has value u, i.e.,
 - ...

```
begin
```

$$\delta^* := \delta(t); \quad x := t;$$

let z (z < u) be the value of the current flow;

if $z + \delta^* > u$ then $\delta^* := u - z$

repeat

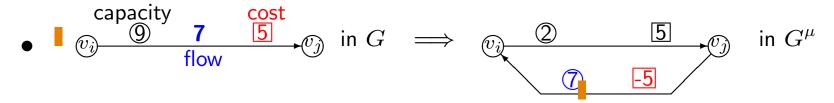
. . .

until x = s;

if z (updated) = u then stop else cancel all labels

The Minimum Cost Flow Problem: Phase 2

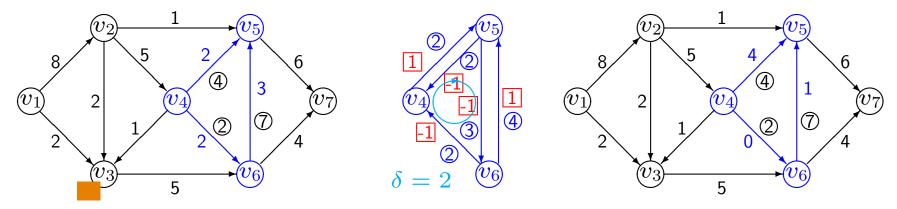
- Given a flow ξ_{ij} $(i,j=1,\ldots,n)$ of value u in G=(V,A), let us define the incremental graph $G^{\mu}=(V,A^{\mu})$ as:
 - $-A^{\mu} = A^{\mu}_f \cup A^{\mu}_b; \blacksquare$
 - $A_f^{\mu} = \{ (v_i, v_j) \in A : \xi_{ij} < q_{ij} \}; \blacksquare$
 - $A_b^{\mu} = \{(v_j, v_i) \in A : \xi_{ij} > 0\};$
 - $\forall (v_i, v_j) \in A_f^\mu$, set capacity $q_{ij}^\mu = q_{ij} \xi_{ij}$ and cost $c_{ij}^\mu = c_{ij}$;
 - $\forall (v_j, v_i) \in A_b^{\mu}$ set capacity $q_{ji}^{\mu} = \xi_{ij}$ and cost $c_{ji}^{\mu} = -c_{ij}$.



- The incremental graph represents the possibilities of increasing/decreasing the current flow.
- Observation:
 - The labeling procedure of the Ford-Fulkerson algorithm can be seen as a method of finding a path in the incremental graph;
 - the augmenting chain is a path where the forward arcs correspond to arcs of A_f^{μ} , and the backward arcs correspond to arcs of A_b^{μ} .

The Minimum Cost Flow Problem: Phase 2 (cont'd)

• Key observation: If the incremental graph G^{μ} contains a circuit Φ such that the sum of the arc costs in Φ is negative then we can send the maximum possible flow δ around the circuit. Example: flow from v_1 to v_7 . Assume all costs are 1.



Flow of value 10 Cost for the blue arcs = 7

In the incremental graph

New flow of value 10 Cost for the blue arcs = 5

- By sending such a flow:
 - the overall flow from the source to the sink remains unchanged and feasible (the equilibrium at each vertex is preserved);
 - the cost is reduced by δ .
- **Property** A flow of value u is a minimum cost flow if and only if the incremental graph G^{μ} contains no circuit such that the sum of the costs in its arcs is negative.

Proof The "only if" part is obvious. The "if" part is omitted.

The Minimum Cost Flow Problem: Algorithm

Observation: the incremental graph can contain arcs with negative cost; a negative cost circuit can be detected by an appropriate $O(n^3)$ shortest path algorithm (modified Dijkstra, or Floyd–Warshall). procedure Minimum_Cost_Flow: begin use a maximum flow algorithm to find a feasible flow $[\xi_{ij}]$ of value u; opt:= false: while opt = false doconstruct the incremental graph G^{μ} corresponding to the current flow $[\xi_{ij}]$; use an appropriate shortest path algorithm to find a negative cost circuit Φ (if any) in G^{μ} ; **if** no such circuit exists **then** opt:= true else begin $\delta := \min_{(v_i, v_j) \in \Phi} \{q_{ij}^{\mu}\}$ for each $(v_i, v_j) \in \Phi$ with $c_{ij}^{\mu} < 0$ do $\xi_{ji} := \xi_{ji} - \delta$; for each $(v_i, v_j) \in \Phi$ with $c_{ij}^{\mu} > 0$ do $\xi_{ij} := \xi_{ij} + \delta$; end endwhile. end.

Complexity of the Minimum Cost Flow Problem

- The algorithm Minimum_Cost_Flow is the most simple and elegant approach, but not the most efficient one.
- Each iteration takes $O(n^3)$ time for detecting a negative circuit;
- the number of iterations can be proportional to $Q = \max_{ij} \{q_{ij}\}$ and $C = \max_{ij} \{c_{ij}\}$, and hence the overall time complexity is pseudo-polynomial:

$$O(n^3 Q C)$$
.

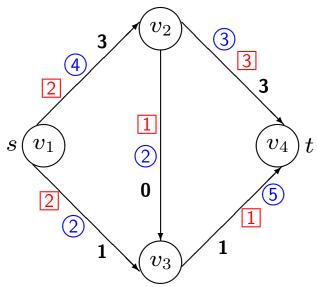
- Other, somehow similar, algorithms are known, that have better pseudo-polynomial time complexity.
- A different family of algorithms, based on scaling techniques, have polynomial time complexity, like, e.g.,

$$O(n^3 \log Q \log(nC));$$

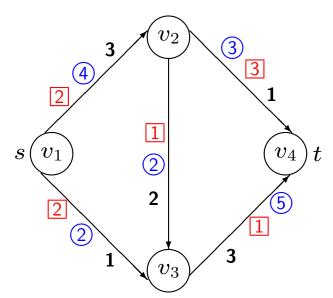
- such time complexities are however called **weakly polynomial**, because they still depend on the magnitude of the input values.
- A third family of (very complicated) algorithms have instead fully polynomial time complexity, like, e.g.,

$$O((n^2 \log n)(n^2 + n \log n)).$$

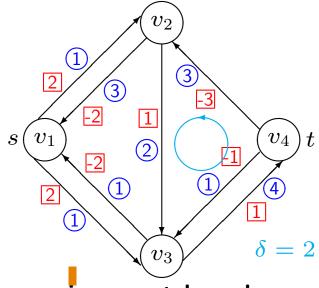
Example



flow of value 4, cost =18

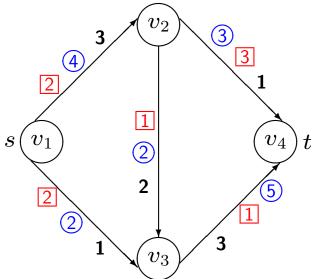


flow of value 4, cost = 16

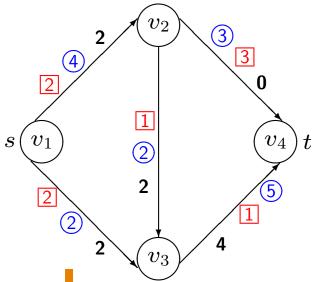


incremental graph

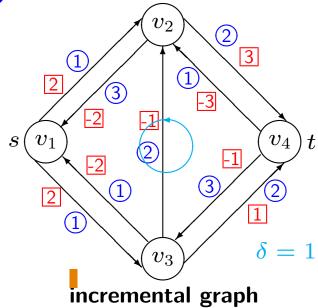
Example (cont'd)

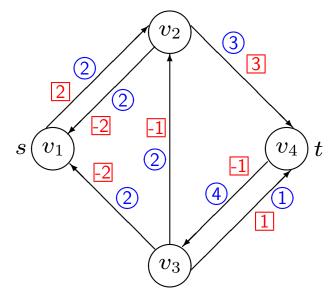


flow of value 4, cost = 16



flow of value 4, cost =14





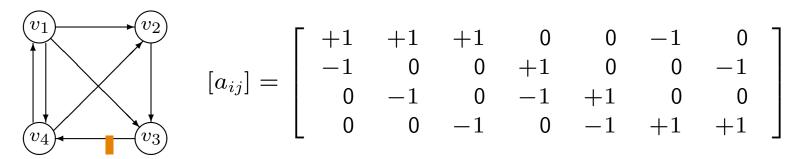
incremental graph no negative cost circuit

Mathematical models

- A special data structure for representing graphs in mathematical models:
- Directed graphs: let a_i (i = 1, ..., m) be the arcs in the graph. Incidence matrix: an $n \times m$ matrix $A = [a_{ij}]$ with

$$a_{ik} = \begin{cases} +1 & \text{if arc } a_k \text{ emanates from vertex } v_i \,; \\ -1 & \text{if arc } a_k \text{ enters vertex } v_i \,; \\ 0 & \text{otherwise} \,. \end{cases}$$

• Example:



Shortest paths and linear programming

- Shortest path from s to t reformulated as:
- send, at minimum cost, a flow of material from s to t:
- $\xi_k =$ quantity of material that flows along arc a_k ;
- c_k (length of arc a_k) = cost to be paid for sending one unit of ξ_k along arc a_k ;
- shortest s to t path = find a minimum cost flow to send one unit of material from s to t.
- row i of A: a "+1" for each arc emanating from v_i , a "-1" for each arc entering v_i :
- flow conservation (β): $a_i'\xi = 0$ for $v_i \neq s, t \Rightarrow$

low conservation
$$(\beta)$$
: $a_i'\xi = 0$ for $v_i \neq s, t \Rightarrow$

$$(P) \min c'\xi \qquad \qquad (D) \max \pi_s - \pi_t \qquad \qquad A_k \qquad \qquad A_k$$

- (P): a constraint per vertex; (D): a constraint per arc, $\pi_i \pi_j \leq c_{ij} = \text{cost of arc } (v_i, v_j)$.
- **Could** ξ be fractional? Yes but: Intuitively, different paths must have the same unit cost. **Formally**, is A a TUM matrix?.

A sufficient unimodularity condition

Theorem An integer matrix A with $a_{ij} \in \{0, +1, -1\} \ \forall \ i, j$ is TUM if

- 1. no column has more than two non-zero elements, and
- 2. the rows can be partitioned into two sets, I_1 and I_2 such that
- if a column has two entries of the same sign, their rows are in different sets;
- if a column has two entries of different sign, their rows are in the same set.

Proof By induction. **Base:** any square submatrix of A of order 1 is UM. **Induction:** we show that:

if every submatrix of order k-1 is UM, then every submatrix of order k is UM.

Let C be a submatrix of order k. Three possibilities exist:

- (a) C has a column of all zero entries: then det(C) = 0;
- (b) C has a column with a single non-zero entry c_{ij} : then $\det(C) = \pm 1 \cdot (\text{minor of } c_{ij})$. The minor of c_{ij} is a submatrix of order k-1, hence UM. It follows that C is UM.
- (c) C all columns of C have two non-zero entries: for each column j we have

$$\sum_{i \in I_1} a_{ij} = \sum_{i \in I_2} a_{ij}.$$

Hence a linear combination of the rows of C has value zero, i.e., $\det(C) = 0$. \Box

Some important consequences

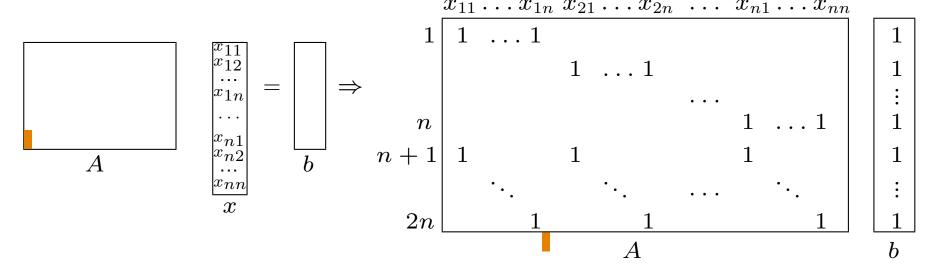
1. The incidence matrix of a directed graph is TUM.

Proof It satisfies the hypothesis with $I_2 = \emptyset \square$.

- 2. The shortest path problem can be solved through linear programming.

 (In addition, it can be proved that the Dijkstra algorithm is equivalent to a specialized version of an algorithm for Linear Programs, known as Primal-Dual.)
- 3. The constraint matrix of the Assignment Problem (AP) is TUM.

Proof Constraints $\sum_{j=1}^{n} x_{ij} = 1 (i = 1, ..., n)$ and $\sum_{i=1}^{n} x_{ij} = 1 (j = 1, ..., n)$ model an LP in standard form with n^2 variables:



which satisfies the hypothesis with $I_1=\{1,\ldots,n\}$, $I_2=\{n+1,\ldots,2n\}$. \square

4. It follows that we can solve the AP by solving through LP its continuous relaxation, obtained by replacing constraints $x_{ij} \in \{0, 1\}$ with $x_{ij} \geq 0$ (constraints $x_{ij} \leq 1$ are redundant).

Maximum flows and linear programming

• Incidence matrix A. Flow of value z from s to t:

$$\xi \leq q$$

$$A\xi = \begin{cases} +z & \text{row } s \\ -z & \text{row } t \\ 0 & \text{rows } \neq s, t \end{cases}$$

$$\xi \geq 0.$$

- The weaker form $A\xi + dz \le 0$ is equivalent to $A\xi + dz = 0$.

Proof For a vertex v_i , the quantity $[a_i'\xi]$ means $[(flow-out \ v_i) - (flow-in \ v_i)]$. Hence:

- (a) row $s: A\xi + dz \le 0 \Rightarrow (\text{flow-out } s) \le z$;
- (b) row t: $A\xi + dz \le 0 \Rightarrow (\text{flow-in } t) \ge z$;
- (c) row $i \neq s, t$: $A\xi + dz \leq 0 \Rightarrow$ (flow-out v_i) \leq (flow-in v_i).
- [(b) and (c)] \Rightarrow (a) (flow-out s) = z;
- [(a) and (c)] \Rightarrow (b) (flow-in t) = z;
- [(b) and (c)] \Rightarrow (c) (flow-out v_i) = (flow-in v_i) $\forall v_i \neq s, t$.

Maximum flows and linear programming (cont'd)

• Resulting LP model in m+1 variables:

$$(D) \max z$$

$$A\xi + dz \leq 0$$

$$\xi \leq q$$

$$-\xi \leq 0$$

$$\xi , z \geq 0 .$$

- (D) can be interpreted as the dual of a primal in standard form.
- \bullet It can be shown that the primal of (D) is equivalent to finding the minimum cut, and that
- the Ford-Fulkerson algorithm is equivalent to a specialized version of a particular LP algorithm, the *Primal-Dual Algorithm*.
- The Primal-Dual Algorithm provides the theoretical foundations of a number of important algorithms for graphs and networks.