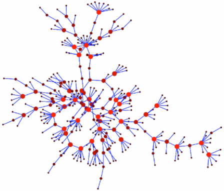
SC42100 – Academic year 2019/2020

Networked and Distributed Control Systems

Introduction and basic graph notions



Giulia Giordano

University of Trento

(Visiting Prof. at TU Delft: `g.giordano@tudelft.nl`)

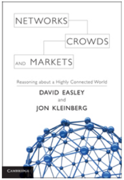
(Special thanks to **Giacomo Como**!)

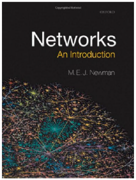
Course Material (First Part: Network Dynamics)

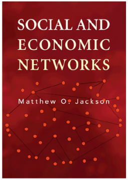
- Slides
  Special Thanks to Giacomo Como!

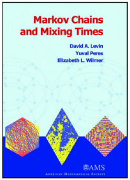
- Lecture notes:
  - *Lecture notes in Network Dynamics*
    by Giacomo Como and Fabio Fagnani
    Material in the slides, more in depth — Reference textbook
  - *Lectures on Network Systems* (Chapters 1-9, available online)
    by Francesco Bullo
    Almost same topics, different approach — Suggested reading

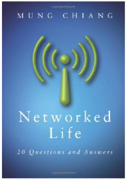
- Several interesting books on the subject, for instance:

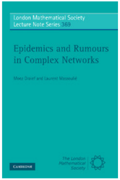












Course contents (First Part: Network Dynamics)

Graphs and Networks: basic notions.
Lecture notes in Network Dynamics, Chapter 1, Sections 1 and 2
Lectures on Network Systems, Chapters 1, 3

Linear algebra on graphs and node centrality.
Lecture notes in Network Dynamics, Chapter 1, Sections 3 and 4
Lectures on Network Systems, Chapters 2, 4, 6; 5.4

Linear network dynamics:
averaging systems and compartmental systems.
Lecture notes in Network Dynamics, Chapter 2
Lectures on Network Systems, Chapters 5, 7, 9

Connectivity and the max-flow min-cut theorem.
Lecture notes in Network Dynamics, Chapter 1, Section 5
Lectures on Network Systems, Chapter 8

Flow optimization.
Lecture notes in Network Dynamics, Chapter 4

(Complex) networks

(Large-scale) systems of (simple?) interacting units

- infrastructure networks: transportation, power, gas, and water distribution; Internet; (wireless) sensor networks
- robotic networks
- informational networks: WWW, citation networks
- social networks: friendships, family ties, Facebook etc.
- economic networks: supply chains, production networks
- biological networks: neural networks, gene/protein interactions
- ecological networks: food webs, flocks, ...

Network dynamics

$$\text{network structure} + \text{interaction mechanism}$$
$$\Downarrow$$
$$\text{emergent behavior}$$

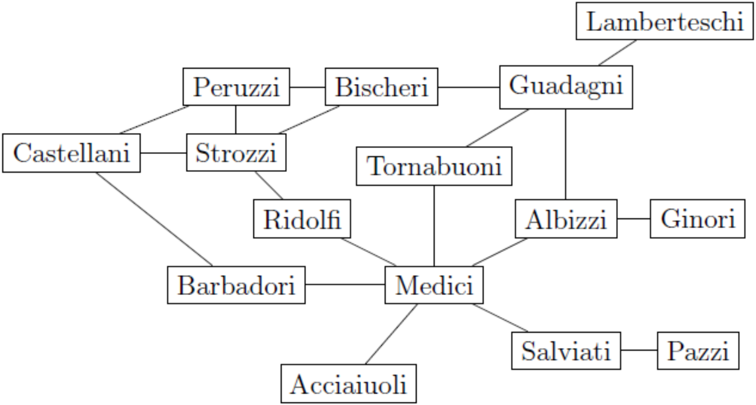
Many applications:

- physical flows in infrastructure networks
- opinion formation and social influence
- design of distributed algorithms
- spread of epidemics and innovation
- cascading failures, systemic risk
- formation control and coordination of multi-robot systems...

Emphasis on common principles:

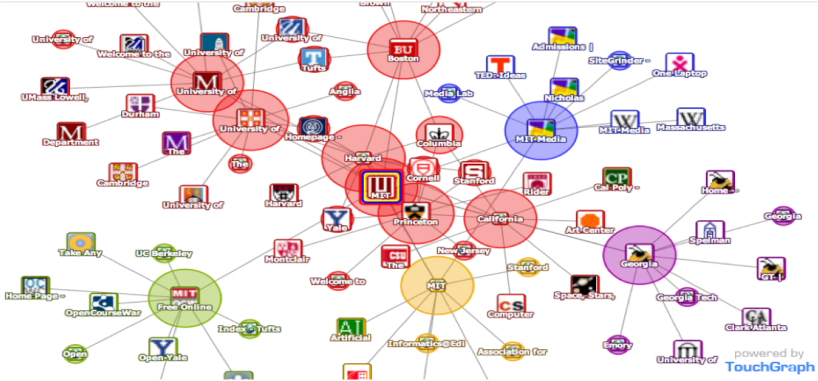
- connection between dynamic evolution and network topology;
- centrality and influence;
- network connectivity, fragility and resilience.

Example 1: Family ties in 15th century Florence



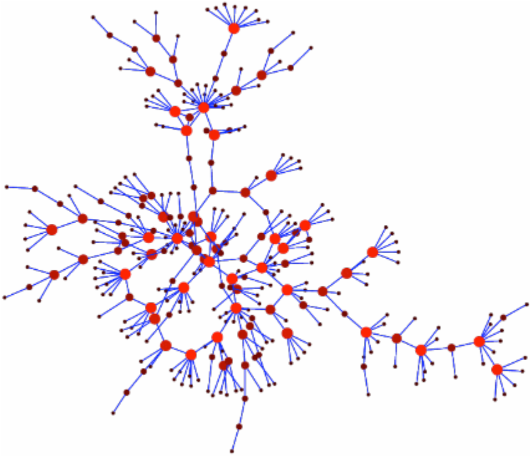
from Padgett and Ansell, 'Robust action and the rise of the Medici, 1400-1434', 1993

Example 2: the World-Wide Web



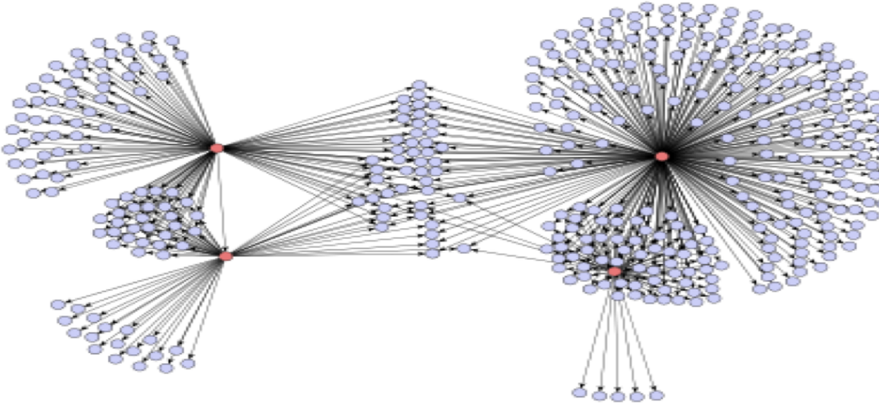
The web link structure centered at http://web.mit.edu

Example 3: Sexual contacts



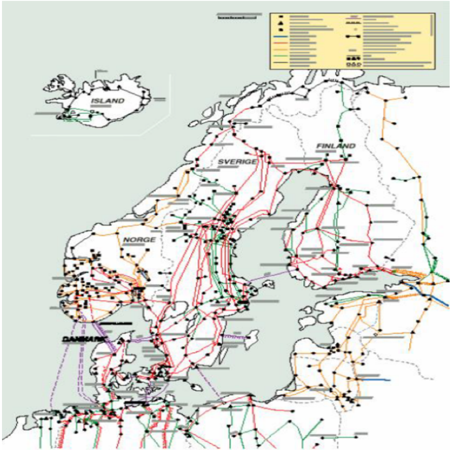
from Potterat et al., ‘Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs’, 2002.

Example 4: Information spread in social network

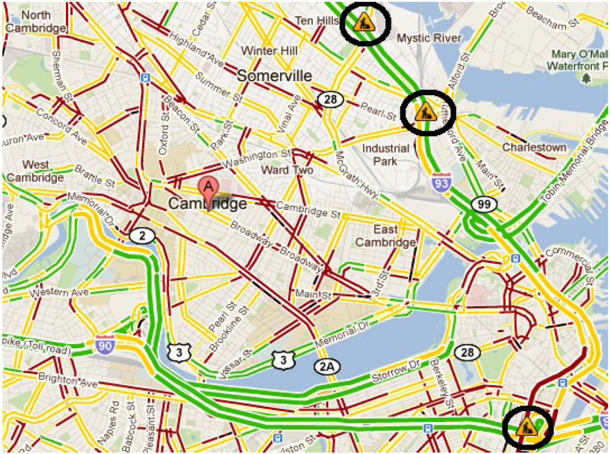


E-mail recommendations for a Japanese graphic novel, from Leskovec 2007.

Example 5: power transmission grid in Scandinavia

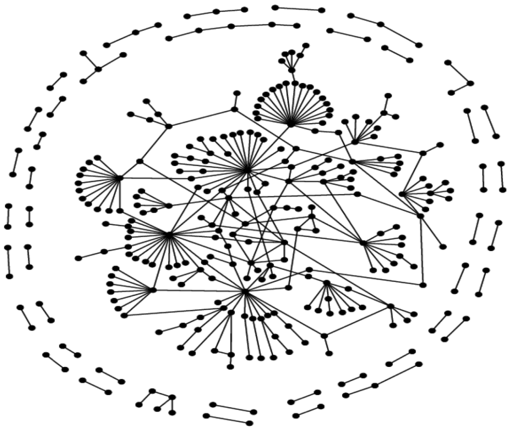


Example 6: urban transportation network



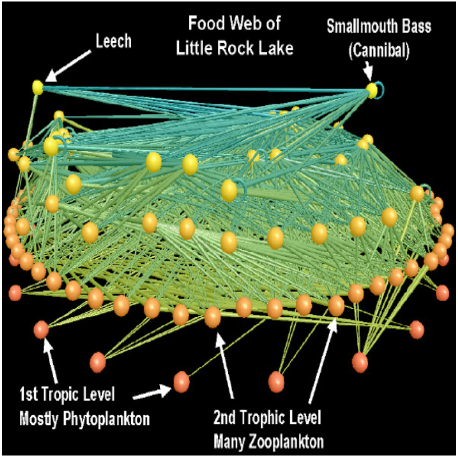
from google maps, Cambridge (MA), July 11, 2011, 18:30 ca.

Example 7: protein network in yeast nucleus



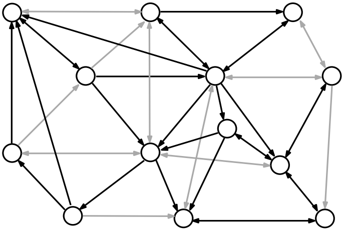
from Maslov and Sneppen 'Specificity and stability in topology of protein networks', 2002

Example 8: Freshwater food web



from Martinez, 'Artifacts or attributes? Effects of resolution on the Little Rock Lake food web', 1991

Networks as graphs

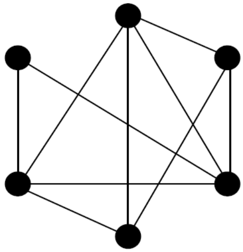


- a graph $\mathcal{G}$ has a set $\mathcal{V} = \{1, \ldots, n\}$ of nodes (or **vertices**) and a set $\mathcal{E}$ of links (or **edges**, arcs)
- links can be directed: (i, j) points from i to j or undirected: $\{i, j\}$ = both (i, j) and (j, i)
- associate to each link a positive weight:

$$\text{weight of link } (i,j) \text{ is } W_{ij} \begin{cases} > 0 & \text{if} \quad (i,j) \in \mathcal{E} \\ = 0 & \text{if} \quad (i,j) \notin \mathcal{E} \end{cases}$$

- weight matrix $W = (W_{ij}) \in \mathbb{R}_{+}^{n \times n}$

Weight matrix / Adjacency matrix

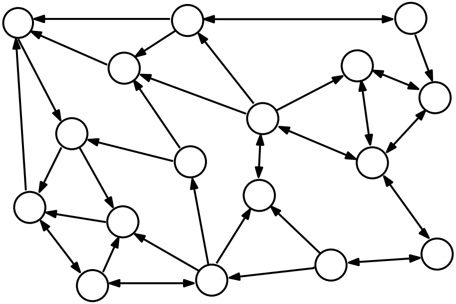


$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

A graph is described by the triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

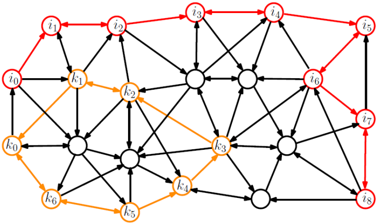
- $\mathcal{G}$ unweighted if $W_{ij} \in \{0, 1\}$ for all i, j; all existing links have weight 1. If so, W is called adjacency matrix and $(\mathcal{V}, \mathcal{E}) \leftrightarrow W$
- $\mathcal{G}$ undirected if $W^\top = W$; all links are bidirectional with the same weight $W_{ij} = W_{ji}$. $\mathcal{G}$ is directed (di-graph) otherwise.
- $\mathcal{G}$ simple if
unweighted $+$ undirected $+$ no self-loops ($W_{ii} = 0$ for all $i \in \mathcal{V}$)

Examples



- Internet: nodes=routers, edges=direct physical links (und.)
- scientific collab.: nodes=researchers, link=coauthors (und.)
- World Wide Web: nodes=webpages, links=hyperlinks (dir., unw.)
- traffic networks: nodes=junctions, links=roads (directed, weighted)

Walks, paths, and cycles



In a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

- a walk from i to j is a sequence of nodes $i = i_0, i_1, \ldots, i_\ell = j$ such that $(i_{h-1}, i_h) \in \mathcal{E}$ for $h = 1, \ldots, \ell$;
- a path from i to j is a walk $i = i_0, i_1, \ldots, i_\ell = j$ that never visits the same node more than once except for possibly $i = j$;
- a cycle is a path that starts and ends in the same node $i = j$ ($i = k_0, k_1, \ldots, k_6, j = k_0$);
- ℓ is called the length of the walk, path, or cycle (# involved links);
- note: length-1 paths = links; length-1 cycles = self-loops;
- $\mathcal{G}$ is called acyclic if it contains no cycles;
- $\mathcal{G}$ is connected if, for all $i, j \in \mathcal{V}$, there exists some path from i to j

Walks, paths, cycles

In a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

- the weight of a length-ℓ walk, path or cycle $i_0, i_1, \ldots, i_\ell$ is $w_{i_0 i_1} \cdot w_{i_1 i_2} \cdot \ldots \cdot w_{i_{\ell-1} i_\ell} = \prod_{1 \le h \le \ell} w_{i_{h-1} i_h}$
- in an unweighted graph all walks, paths and cycles have weight 1

Proposition: In an unweighted graph,

$$\#\{\text{length-}\ell \text{ walks from } i \text{ to } j\} = (W^\ell)_{ij}.$$

For a general weighted graph

$$\text{total weight of all length-}\ell \text{ walks from } i \text{ to } j = (W^\ell)_{ij}.$$

Distance and diameter

- distance of j from i:

$$\mathrm{dist}(i,j) = \text{length of shortest path from } i \text{ to } j$$

- diameter of graph $\mathcal{G}$:

$$\mathrm{diam}(\mathcal{G}) = \max_{i,j\in\mathcal{G}} \mathrm{dist}(i,j)$$

- $\mathcal{G}$ connected if and only if $\mathrm{diam}(\mathcal{G}) < +\infty$

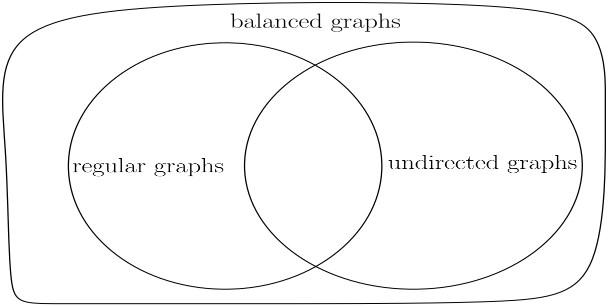
Neighborhoods and degrees

- out-neighborhood of node i: $\mathcal{N}_i = \{j : (i,j) \in \mathcal{E}\}$
- in-neighborhood of node i: $\mathcal{N}_i^- = \{j : (j,i) \in \mathcal{E}\}$
- $\mathcal{G}$ undirected $\Longrightarrow \mathcal{N}_i^- = \mathcal{N}_i$ for all i
- out-degree $w_i = \sum_{j \in \mathcal{V}} W_{ij}$ (row sum)
- in-degree $w_i^- = \sum_{j \in \mathcal{V}} W_{ji}$ (column sum)
- degree vectors: $w = W\mathbb{1}$ and $w^- = W^\top \mathbb{1}$
- average degree:

$$\overline{w} = \frac{1}{n}\mathbb{1}^\top W \mathbb{1} = \frac{1}{n}\sum_i w_i = \frac{1}{n}\sum_i w_i^-$$

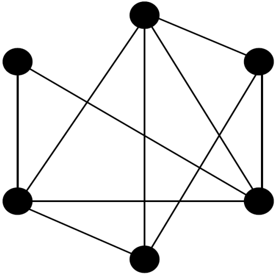
- $\mathcal{G}$ unweighted $\Longrightarrow$ $w_i = |\mathcal{N}_i|$ $\quad w_i^- = |\mathcal{N}_i^-|$

Balanced, regular, and undirected graphs



- a graph is called balanced if $w^{-} = w$
- clearly every undirected graph is balanced, but not vice versa
- a graph is called regular if $w_i = w_i^{-} = \overline{w}$ for all i
- every regular graph is balanced, but not vice versa; neither regular graphs are necessarily undirected, nor vice versa.

Average degree, hand-shaking lemma

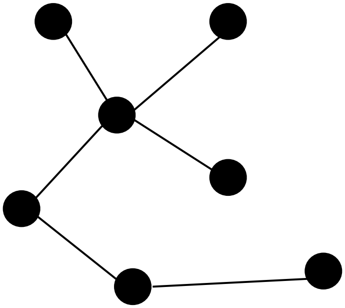


$$4 + 3 + 4 + 3 + 4 + 2 = 2 \cdot 10$$

In a simple graph the total degree is equal to twice the number of undirected links

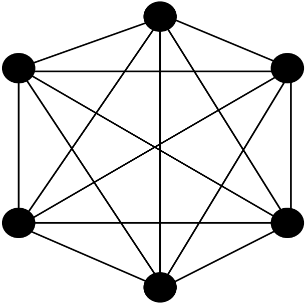
$$n\overline{w} = \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} w_{ij} = 2m$$

Trees



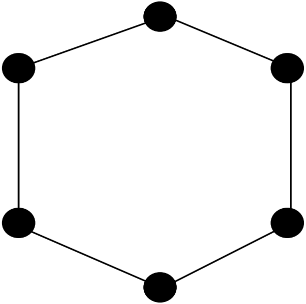
- a tree is a simple connected graph with $m = n - 1$ undirected links
- degree-1 nodes in a tree are called leaves
- the number of leaves of a tree is ≥ 2 and $\leq n - 1$ (prove it!)
- a tree with 2 leaves is called a line
- a tree with $n - 1$ leaves is called a star

Other special simple graphs



complete graph

$m = n(n-1)/2$

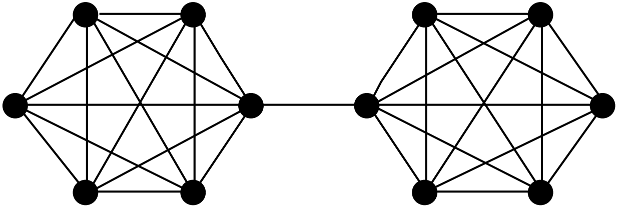


ring graph

$m = n$

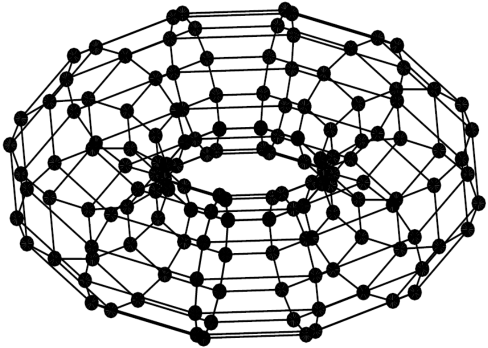
Other special simple graphs

- barbell:



$$m = \frac{n}{2}\left(\frac{n}{2} - 1\right) + 1$$

- torus: $m = 2n$



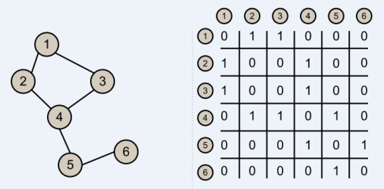
What we learned today

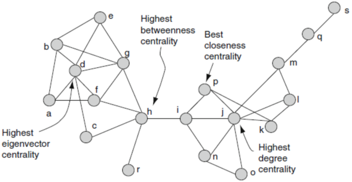
- Why networks and networked dynamics are relevant and interesting
- How to model networks as graphs
- Basic notions: adjacency and weight matrices; walks, paths and cycles; degree distributions; distance and diameter; clustering and modularity.
- “Food-for-thought” application: the structure of Facebook.

SC42100 – Academic year 2019/2020

Networked and Distributed Control Systems

Algebraic graph theory and node centrality





Giulia Giordano

University of Trento

(Visiting Prof. at TU Delft: `g.giordano@tudelft.nl`)

(Special thanks to **Giacomo Como**!)

Matrices associated with a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

- Weight/adjacency matrix: $W \in \mathbb{R}_+^{n \times n}$,

$$\text{weight of link } (i,j) = W_{ij} \begin{cases} > 0 & \text{if} \quad (i,j) \in \mathcal{E} \\ = 0 & \text{if} \quad (i,j) \notin \mathcal{E} \end{cases}$$

.

W is a nonnegative matrix!

Assume $w_i = \sum_{j \in \mathcal{V}} W_{ij} > 0$ for all i
(if $w_i = 0$ for some i, add self-loop on i with $W_{ii} > 0$).

Recall $w = W\mathbb{1}$.

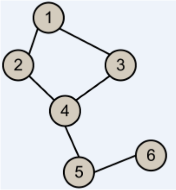
- Degree matrix: $D = \operatorname{diag}(w) \in \mathbb{R}_+^{n \times n}$

- Laplacian matrix: $L = D - W \in \mathbb{R}^{n \times n}$

- Normalized weight/adjacency matrix: $P = D^{-1}W \in \mathbb{R}_+^{n \times n}$

P is a nonnegative matrix!

Matrices associated with a graph: example



$$W = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad w = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 3 \\ 2 \\ 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix},$$

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Laplacian and normalized adjacency matrix: properties

$w = W\mathbb{1}, \qquad D = \operatorname{diag}(w)$

- $L = D - W \in \mathbb{R}^{n \times n}$

  $L\mathbb{1} = w - w = 0 \Longrightarrow 0$ is an eigenvalue of L

  $-L$ is a Metzler matrix (nonnegative off-diagonal entries)

Undirected graph: W and L are symmetric $\Longrightarrow$ all eigenvalues real, $\exists$ an orthonormal basis of $\mathbb{R}^n$ formed by eigenvectors

- $P = D^{-1}W \in \mathbb{R}_+^{n \times n}$

  $P\mathbb{1} = \mathbb{1} \Longrightarrow 1$ is an eigenvalue of P

P is a stochastic matrix: nonnegative, square, rows sum up to one

Undirected graph: P similar to a symmetric matrix $D^{-1/2}WD^{-1/2}$ $\Longrightarrow$ same spectral properties

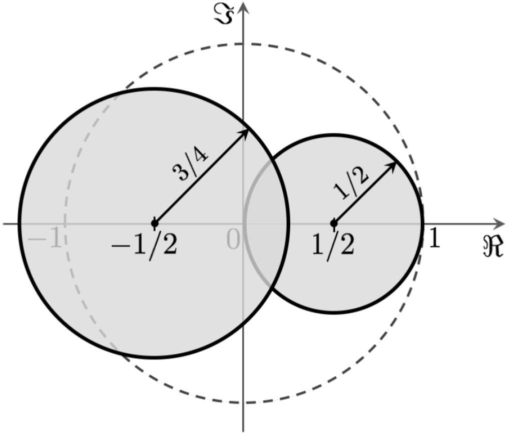
Undirected and regular graph: P symmetric $\rightarrow$ doubly stochastic, both columns and rows sum up to one!

Gershgorin theorem

In the complex plane, the eigenvalues of matrix A lie in the union of the n disks with center a_{ii} and radius $\sum_{j \neq i} |a_{ij}|$, for $i = 1, \ldots, n$.

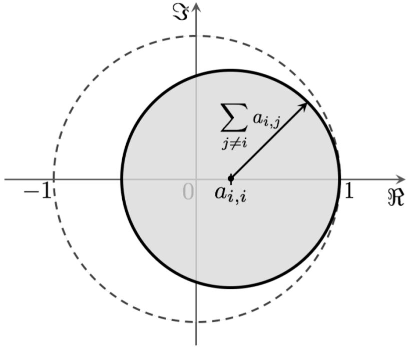
Example

$$A = \begin{bmatrix} 1/2 & -1/2 \\ 3/4 & -1/2 \end{bmatrix}$$



Gershgorin theorem for graphs

If A stochastic, $\sum_{j=1}^{n} a_{ij} = 1$, namely $\sum_{j \neq i} a_{ij} = 1 - a_{ii}$.



This applies to matrix P!

Laplacian matrix L: $\sum_{j=1}^{n} L_{ij} = 0$, non-negative diagonal entries and non-positive off-diagonal entries, namely $\sum_{j \neq i} |L_{ij}| = L_{ii}$, so all disks in the right half-plane, touching 0

Perron-Frobenius theory for nonnegative matrices

Mathematicians Oskar Perron (1907) and Georg Frobenius (1912).

Perron-Frobenius Theorem

Let $M \in \mathbb{R}_+^{n \times n}$ be a nonnegative matrix. Then,

- M admits an eigenvalue λ_M whose modulus satisfies $|\lambda_M| \geq |\mu|$ for all other eigenvalues μ of M
- λ_M is a nonnegative real number
- λ_M is associated with left and right eigenvectors having nonnegative entries.

If M is also irreducible ($\sum_{k=0}^{n-1} M^k$ is positive; strongly connected graph), then the eigenvalue λ_M is strictly positive and simple and the left and right eigenvectors have positive entries and are unique, up to rescaling.

If M is also primitive (M^k is positive for some $k \in \mathbb{N}$; strongly connected and aperiodic graph), then $\lambda_M > |\mu|$ for all other eigenvalues μ of M.

Implications for graphs

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, Laplacian L, normalized adjacency matrix P.

(i) The largest-in-modulus eigenvalue λ of W and of $W^\top$ is real and nonnegative, and there exist nonnegative eigenvectors x and y such that $Wx = \lambda x$ and $W^\top y = \lambda y$

(ii) The largest-in-modulus eigenvalue of P and $P^\top$ is 1

(iii) 0 is an eigenvalue of L and $L^\top$ and all their other eigenvalues have positive real part

(iv) w is an invariant vector of $P^\top$ if and only if $\mathcal{G}$ is balanced

(v) $\mathbb{1}$ is an invariant vector of $P^\top$ if $\mathcal{G}$ is regular

(vi) $\mathbb{1}$ is in the kernel of $L^\top$ if and only if $\mathcal{G}$ is balanced

(vii) -1 is an eigenvalue of P and $P^\top$ if and only if $\mathcal{G}$ is bipartite

(viii) If $\mathcal{G}$ is connected, then the geometric multiplicities of λ as an eigenvalues of W, of 1 as an eigenvalue of P, and of 0 as an eigenvalue of L are all equal to 1

Connected components

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ (strongly) connected if, for every two nodes $i, j \in \mathcal{V}$, there exists a path from i to j with finite length l (equivalently, $\mathrm{diam}(\mathcal{G})$ is finite).

Connected components of $\mathcal{G}$: maximal subsets $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k$ of $\mathcal{V}$ such that, for every two nodes i and j in the same component $\mathcal{V}_h$, there exists a path from i to j.

Connected components are a partition of the node set $\mathcal{V}$:
$\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2 \cup \ldots \mathcal{V}_k$ and $\mathcal{V}_h \cap \mathcal{V}_l = \emptyset$ for all $h \neq l$.
Every node belongs to one and only one connected component.

Connected components and algebraic properties: undirected graph

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ undirected graph $\Longrightarrow$ there are no links between nodes belonging to different connected components.

$c_{\mathcal{G}}$: number of connected components.

Proposition The (algebraic and geometric) multiplicities of 0 as an eigenvalue of L and of 1 as an eigenvalue of P are $c_{\mathcal{G}}$.

Connected components and algebraic properties: directed graph

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ directed graph $\Longrightarrow$ there could be links pointing from nodes in a connected component towards nodes in another connected component but not vice versa.

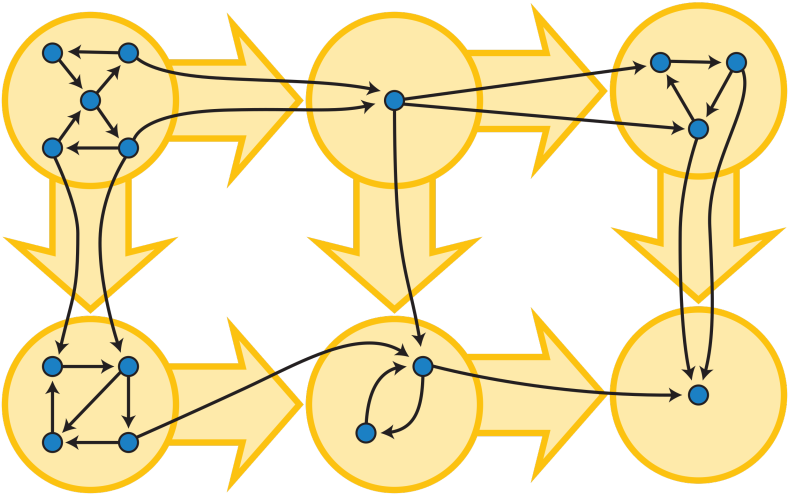
Condensation graph $\mathcal{H}$: nodes in every connected component of $\mathcal{G}$ collapsed into single 'supernodes' of $\mathcal{H}$, link in $\mathcal{H}$ from supernode i to i when there is a link in $\mathcal{G}$ pointing from a node in the connected component i to a node in the connected component j.

$\mathcal{H}$ is acyclic; is a single node iff $\mathcal{G}$ is connected; has a single sink iff there exists one node in $\mathcal{G}$ that is reachable from every other node.

$s_{\mathcal{G}} \geq 1$: number of sinks of $\mathcal{H}$.

Proposition The geometric multiplicities of 0 as an eigenvalue of L and of 1 as an eigenvalue of P are $s_{\mathcal{G}}$.

Condensation graph (directed graphs)



Connected components are in the yellow circles ('supernodes').

Top-left 'supernode': **source**

Bottom-right 'supernode': **sink**

Node centrality: different measures

Centrality measures: importance of a node in graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ depends from its context (surroundings, topology)

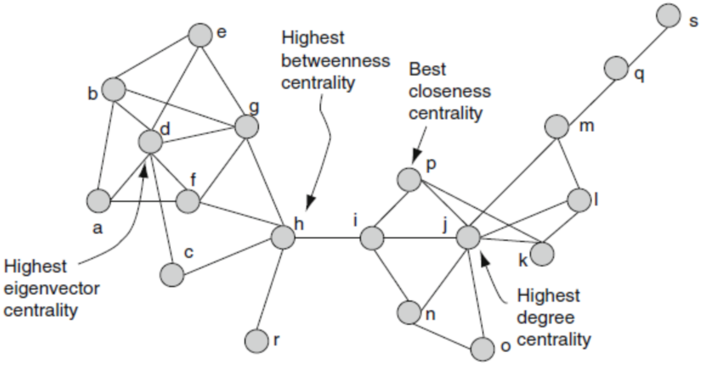
- Degree centrality

- Eigenvector centrality (unnormalized, no intrinsic centrality)
- Bonacich centrality (normalized, no intrinsic centrality)
- Katz centrality (unnormalized, with intrinsic centrality)
- PageRank centrality (normalized, with intrinsic centrality)

- Hub and Authority centrality
- Closeness and Betweenness

Node centrality: different measures

Different measures of node centrality, may not coincide



Degree centrality

Degree centrality: importance of node i is its degree.

- Out-degree (# links originating from i)?
- In-degree (# links pointing to i)?
- Typically in-degree (e.g., followers, citations)

Extensions of in-degree centrality

Centrality π_i of a node i increases
more due to connections from nodes with high centrality
than due to connections from nodes with low centrality

Centrality of a node proportional to sum of centralities of its in-neighbors

$$\pi_i \propto \sum_j W_{ji} \pi_j$$

$\pi \in \mathbb{R}^n$ vector of node centralities (centrality vector), $\pi \propto W^\top \pi$

Eigenvector centrality

$\lambda\pi = W^\top \pi$: proportionality constant $1/\lambda > 0$
π is an eigenvector of $W^\top$ corresponding to a positive eigenvalue λ

Recall: the largest-in-modulus eigenvalue λ^* of $W^\top$ is positive and real, and is associated with a nonnegative eigenvector

Eigenvector centrality: nonnegative eigenvector of $W^\top$ corresponding to the largest-in-modulus positive real eigenvalue λ^*:

$$\lambda^* \pi = W^\top \pi$$

- π not unique: any rescaling is also an eigenvector of $W^\top$.
Total centrality of all nodes fixed to a given number
(n or, better, 1, so that π is a probability vector)

- Just for connected graphs (eigenvector $\pi = (\lambda^*)^{-1} W^\top \pi$ unique up to rescaling).
For not connected graphs, there could be linearly independent nonnegative eigenvectors of W associated with λ^*

Bonacich centrality

We do not want nodes to contribute to the centrality of all their out-neighbors in the same way, regardless of their out-degree

Bonacich centrality: normalize the out-degree as $\pi_i \propto \sum_j \frac{W_{ji}}{w_j}\pi_j$, i.e., replace W with $P = D^{-1}W$
($D = \mathrm{diag}(w)$ and, if $w_i = 0$, $D_{ii} = 1$ and $D_{ij} = 0$ for all $j \neq i$).

P stochastic matrix, largest-in-modulus eigenvalue is 1:

$$\pi = P^\top \pi$$

The Bonacich centrality π is an invariant vector of $P^\top$.

Directed graphs: the number of linearly independent invariant vectors of $P^\top$ equals $s_{\mathcal{G}}$.

Eigenvector and Bonacich centrality: drawbacks

- Add self-loop of very large weight on node i:
weight $\to \infty \implies$ ratio(centrality of i/total centrality of all other nodes) $\to \infty$

- Add undirected link of very large weight between two nodes:
weight $\to \infty \implies$ ratio(sum of the two nodes' centralities/total the centralities of all other nodes) $\to \infty$

What shall we do?

Nodes must get some centrality independent of their in-neighbors!

Katz centrality

intrinsic centrality μ, nonnegative vector,
standard choice $\mu = \mathbb{1}$ (intrinsic centrality equal for all nodes)

λ dominant eigenvalue of $W^\top$

$\beta \in (0, 1]$ weight of intrinsic centrality relative to network topology

Katz centrality: $\pi = \left(\frac{1-\beta}{\lambda}\right) W^\top \pi + \beta\mu$

$0 < \beta \leq 1 \implies$ dominant eigenvalue of $\lambda^{-1}(1-\beta)W^\top$ is $< 1 \implies$ $[I - \lambda^{-1}(1-\beta)W^\top]$ invertible

$$\pi = \left[I - \lambda^{-1}(1-\beta)W^\top\right]^{-1} \beta\mu = \beta \sum_{k \geq 0} \lambda^{-k}(1-\beta)^k (W^\top)^k \mu$$

If $\beta \to 0$, then Katz centrality $\to$ Eigenvector centrality

For $\beta = 1$, $\pi = \mu$: Katz centrality independent of graph structure

PageRank centrality

Normalized: use P instead of W

PageRank centrality: solution of

$$\pi = (1-\beta)\, P^\top \pi + \beta\mu$$

Relative importance of webpages (Brin & Page); typically $\beta \approx 0.15$

Dominant eigenvalue of $(1-\beta)\, P^\top$ is $1-\beta$, with $0 < \beta \leq 1$
$\Longrightarrow \left[I - (1-\beta)P^\top\right]$ invertible

$$\begin{aligned}\pi &= \left(I - (1-\beta)P^\top\right)^{-1} \beta\mu = \beta \sum_{k\geq 0} (1-\beta)^k (P^\top)^k \mu \\ &= \beta\mu + \beta(1-\beta)P^\top\mu + \beta(1-\beta)^2(P^\top)^2\mu + \dots\end{aligned}$$

Hub centrality and Authority centrality

Hub centrality of a node is proportional to the sum of the authority centralities of its out-neighbors: $x_i \propto \sum_j W_{ij} y_j$

$$x = \alpha W y, \qquad \alpha > 0$$

Authority centrality of a node is proportional to the sum of the hub centralities of its in-neighbors: $y_i \propto \sum_j W_{ji} x_j$

$$y = \beta W^\top x \qquad \beta > 0$$

Hence,

$$\lambda x = W W^\top x\,, \qquad \lambda y = W^\top W y\,,$$

where $\lambda = (\alpha\beta)^{-1}$ is the dominant eigenvalue of $WW^\top$ and $W^\top W$

Note: hub centrality x (authority centrality y) is the eigenvector centrality of the graph with weight matrix $W^\top W$ ($WW^\top$)

Closeness and Betweenness

Closeness centrality: inverse average distance from i to all other nodes j in $\mathcal{V}$

$$\text{closeness}(i) = \frac{n}{\sum_{j \in \mathcal{V}} \text{dist}(i, j)}, \qquad i \in \mathcal{V}$$

Betweenness centrality: for every two nodes i and j, consider all minimum-distance paths from i to j (if any) and let $g_{ij}^{(k)}$ be the fraction of such paths passing through node k

$$\text{betweenness}(k) = \frac{1}{n^2} \sum_{i,j \in \mathcal{V}} g_{ij}^{(k)}, \qquad k \in \mathcal{V}$$

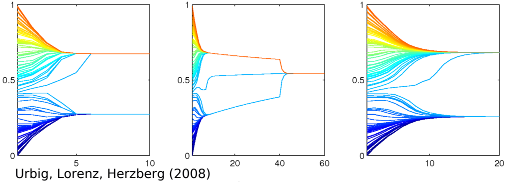
What we learned today

- Algebraic graph theory: matrices associated with a graph, adjacency matrix and Laplacian matrix
- Perron-Frobenius theory for nonnegative matrices, and consequent properties of adjacency and Laplacian matrices
- Connected components
- Different measures for node centrality: how "important" is a node?

SC42100 – Academic year 2019/2020

Networked and Distributed Control Systems

Linear network dynamics: averaging systems and compartmental systems



Giulia Giordano

University of Trento

(Visiting Prof. at TU Delft: `g.giordano@tudelft.nl`)

(Special thanks to **Giacomo Como**!)

Linear network dynamics

Two dual deterministic linear network dynamics

- distributed averaging
(opinion dynamics, distributed estimation in sensor networks, rendezvous problems in mobile robotics networks)

$\rightarrow$ nodes are agents updating their states to a weighted average of their neighbor states

- linear compartmental systems
(dynamical flow networks in biology or infrastructure networks)

$\rightarrow$ nodes are units exchanging flows of some commodity

Distributed averaging (discrete-time): graph

Opinion dynamics in social networks $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$

- nodes in $\mathcal{V}$: population of individuals / agents
- links in $\mathcal{E}$: interactions among agents
- W_{ij}: strength of the influence that agent j has on agent i

- link $(i, j) \iff$ agent i follows (is influenced by) agent j

- Any sink of $\mathcal{G}$ has a self-loop with positive weight: connectivity of $\mathcal{G}$ unchanged, out-degree w_i strictly positive for all nodes

- Agent i has out-degree $w_i = \sum_{j \in \mathcal{V}} W_{ij}$ and state $x_i(t) \in \mathbb{R}$: state is agent opinion, updated at discrete time steps $t = 0, 1, \ldots$ in response to the current states of the out-neighbors

Distributed averaging (discrete-time): dynamics

New opinion $x_i(t+1)$ of agent i: weighted average of the current opinions $x_j(t)$ of the out-neighbors $j \in \mathcal{N}_i$, with weight $\propto W_{ij}$

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i} \frac{W_{ij}}{w_i} x_j(t)\,, \qquad i \in \mathcal{V}$$

Self-loops in the graph $\mathcal{G}$ (nonzero diagonal entries of W): inertia in the update rule, W_{ii}/w_i measure self-confidence of agent i

Recall: $P = D^{-1}W$, $D = \mathrm{diag}\,(w)$, $w = W\mathbb{1}$

- $x(t) = (x_i(t))_{i \in \mathcal{V}}$: vector with opinions of all agents

Compact form: DeGroot's model of opinion dynamics

$$x(t+1) = Px(t)\,, \qquad t = 0, 1, \ldots$$

DeGroot's model of opinion dynamics: equilibria

Equilibrium vector $\bar{x} = P\bar{x}$: eigenvector of the stochastic matrix P associated with largest-in-modulus eigenvalue 1

$P\mathbb{1} = \mathbb{1}$: a possible solution is $\bar{x} \propto \mathbb{1}$, consensus vectors (all entries are equal)

Every consensus vector is equilibrium for DeGroot opinion dynamics

If $\mathcal{G}$ connected, all possible eigenvectors are $\bar{x} \propto \mathbb{1}$: the only equilibria are consensus vectors

(In general, the space of equilibria has dimension equal to the number of sinks $s_{\mathcal{G}}$ of the condensation graph $\mathcal{H}$ of $\mathcal{G}$)

Assume $\mathcal{G}$ connected and study asymptotic behavior:
(1) does opinion dynamics always converge to a consensus $\bar{x}\mathbb{1}$?
(2) how does $\bar{x}$ depend on the initial opinion $x(0)$?

DeGroot's model of opinion dynamics: convergence

Can we guarantee that the opinion vector $x(t)$ actually converges?

Connectivity of $\mathcal{G}$ is not enough!

Example: network with two nodes connected by one undirected link of weight 1, no self loops

$$P = W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Unless already started at a consensus, the dynamics keep on oscillating (both nodes copy one the other's opinion)

These pathological situations can be ruled out if graph $\mathcal{G}$ is connected and aperiodic (maximum common divisor of all cycle lengths is 1; true whenever there is at least one self-loop, i.e., at least one diagonal entry of the weight matrix W is strictly positive)

Self-loops denote agents with self-confidence: natural assumption

DeGroot's model of opinion dynamics: consensus value

$P^\top$ has largest-in-modulus eigenvalue 1, associated with nonnegative eigenvector $\pi = P^\top \pi$: Bonacich centrality vector

$\mathcal{G}$ connected $\implies$ π unique up to a multiplicative constant
normalize π as a probability vector, $\sum_{i \in \mathcal{V}} \pi_i = 1$

$$\pi^\top x(t+1) = \pi^\top P x(t) = (P^\top \pi)^\top x(t) = \pi^\top x(t)\,, \qquad t = 0, 1, \dots\,,$$

the weighted average $\pi^\top x(t)$ is invariant for the dynamics:

$$\pi^\top x(t) = \pi^\top x(0) = \sum_{i \in \mathcal{V}} \pi_i x_i(0)\,, \qquad t = 0, 1, \dots\,.$$

$\mathcal{G}$ connected $\implies$ the only equilibria are consensus vectors: if the opinion vector $x(t)$ converges, its limit for $t \to \infty$ is $x(\infty) = \bar{x}\mathbb{1}$ with consensus value $\bar{x} = \pi^\top x(0)$: weighted average of agent initial opinions $x_i(0)$, weighted by Bonacich centrality π_i

For balanced graphs (hence for undirected graphs) node centrality is normalized degree, $\pi_i = w_i/(n\bar{w})$

DeGroot's model of opinion dynamics: theorem

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ connected aperiodic graph
$P = D^{-1}W$, where $D = \mathrm{diag}\,(w)$, normalized weight matrix.
For every initial opinion vector $x(0) \in \mathbb{R}^n$, the DeGroot opinion dynamics (discrete-time distributed averaging dynamics)

$$x(t+1) = Px(t)\,, \qquad t = 0, 1, \ldots$$

converges to a consensus vector

$$\lim_{t \to +\infty} x(t) = \bar{x}\mathbb{1}\,,$$

with consensus value

$$\bar{x} = \pi^\top x(0) = \sum_{i \in \mathcal{V}} \pi_i x_i(0)\,,$$

where $\pi = P^\top \pi$ is the normalized Bonacich centrality vector.

Application: social learning and wisdom of crowds

Opinion dynamics: social aggregation of information that produces a common final estimate $\bar{x}$ of the state of the world $\theta \in \mathbb{R}$

Initial opinion of each agent $i \in \mathcal{V}$: a noisy version of θ

$$x_i(0) = \theta + \xi_i\,, \qquad i \in \mathcal{V}\,,$$

with ξ_i random variables (noise) with zero mean: expectation of $x_i(0)$ and of $\bar{x}$ is θ (unbiased estimators)

If $\bar{x}$ is a better estimate of θ than any $x_i(t)$, wisdom of crowd

Independent noises ξ_i with variances σ_i^2: variance of the asymptotic consensus value $\bar{x}$ satisfies $\sigma_{\bar{x}}^2 = \sum_i \pi_i^2 \sigma_i^2$
If all $\sigma_i^2 = \sigma^2$, then

$$\sigma_{\bar{x}}^2 = \sigma^2 \sum_i \pi_i^2$$

Since $\pi_i < 1$ for all i, $\sum_i \pi_i^2 < \sum_i \pi_i = 1$, hence $\sigma_{\bar{x}}^2 < \sigma^2$: crowd wiser than any single individual!

Application: distributed averaging in sensor networks 1

$\mathcal{V}$ set of sensors that collect measurements; limited communication & computation capabilities; information exchange with close enough sensors only: graph $\mathcal{G}$ has undirected link between node i and node j if they can communicate

$x_i(0)$ measurement of node $i \in \mathcal{V}$, vector $x(0) \in \mathbb{R}^n$

Iterative distributed algorithm to average sensor measurements

$$x^* = \frac{1}{n} \sum_{i \in \mathcal{V}} x_i(0) = \frac{1}{n} \mathbb{1}^\top x(0)$$

nodes update their state based on information from neighbors only

$x(t+1) = Px(t)$: converge to consensus vector $\lim_{t \to \infty} x(t) = \bar{x}\mathbb{1}$ with consensus value $\bar{x} = \sum_{i \in \mathcal{V}} \pi_i x_i(0)$

If $\mathcal{G}$ regular, $\pi_i = 1/n$ for all i: average!

Otherwise, $\bar{x}$ is not the arithmetic average $\frac{1}{n} \sum_{i \in \mathcal{V}} x_i(0)$, but a weighted version $\sum_{i \in \mathcal{V}} \pi_i x_i(0)$.

Application: distributed averaging in sensor networks 2

Graph is undirected, hence balanced: $\pi_i = w_i/(n\bar{w})$.

Every sensor knows its degree w_i: normalize the initial condition

$$y_i(0) = \frac{x_i(0)}{w_i}\,, \qquad i \in \mathcal{V}\,, \qquad y(t+1) = Py(t)$$

Average of initial measurements divided by average degree $\bar{w}$:

$$\lim_{t\to+\infty} y_i(t) = \sum_j \pi_j y_j(0) = \sum_j \frac{w_j}{\bar{w}n}\frac{x_j(0)}{w_j} = \frac{1}{\bar{w}}\left(\frac{1}{n}\sum_j x_j(0)\right)$$

Distributed consensus algorithm computes the average degree.

$$z(t+1) = Pz(t)\,, \qquad z_i(0) = \frac{1}{w_i}\,, \qquad i \in \mathcal{V}$$

$$\lim_{t\to+\infty} z_i(t) = \sum_j \pi_j z_j(0) = \sum_j \frac{w_j}{\bar{w}n}\frac{1}{w_j} = \frac{1}{\bar{w}}$$

Combining the two distributed consensus algorithms,

$$\lim_{t\to+\infty} \frac{y_i(t)}{z_i(t)} = \frac{x^*/\bar{w}}{1/\bar{w}}$$

Continuous-time distributed averaging

Continuous-time version of distributed averaging $x(t+1) = Px(t)$:

$$\dot{x}_i = \sum_j W_{ij}(x_j - x_i)\,, \qquad i \in \mathcal{V} \,: \qquad \dot{x} = -Lx$$

where $L = D - W$, with $D = \mathrm{diag}\,(w)$, is the graph Laplacian.

Since $L\mathbb{1} = 0$, consensus vectors $x = \bar{x}\mathbb{1}$ are equilibria. If the graph is connected, the only equilibria are consensus vectors.

$$L^\top D^{-1}\pi = (D - W^\top)D^{-1}\pi = \pi - (D^{-1}W)^\top \pi = \pi - P^\top \pi = 0$$

Nonnegative vector $\bar{\pi} = D^{-1}\pi / (\mathbb{1}^\top D^{-1}\pi)$ with entries

$$\bar{\pi}_i = \frac{\pi_i / w_i}{\sum_j \pi_j / w_j}$$

Then, the quantity $\sum_i \bar{\pi}_i x_i(t)$ remains constant in time:

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_i \bar{\pi}_i x_i(t) = \bar{\pi}^\top \dot{x} = \bar{\pi}^\top L x = (L^\top \bar{\pi})^\top x = 0$$

Continuous-time distributed averaging: theorem

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ connected graph,
$D = \mathrm{diag}\,(w)$, $L = W - D$ Laplacian matrix.

For every initial opinion vector $x(0) \in \mathbb{R}^n$, the continuous-time distributed averaging dynamics

$$\dot{x} = -Lx$$

converges to a consensus vector

$$\lim_{t \to +\infty} x(t) = \bar{x}\mathbb{1}\,,$$

with consensus value

$$\bar{x} = \bar{\pi}^\top x(0)$$

where $\bar{\pi} = D^{-1}\pi/(\mathbb{1}^\top D^{-1}\pi)$ and $\pi = P^\top \pi$ is the normalized Bonacich centrality vector.

Aperiodicity is not needed! If graph connected, continuous-time distributed averaging dynamics converge to a consensus vector.

Distributed averaging over non-connected networks: opinion dynamics with stubborn nodes

$\mathcal{G}$ not connected: algebraic and geometric multiplicity of 1 as an eigenvalue of P and $P^\top$ is $s_{\mathcal{G}}$, # sinks in the condensation graph.

Set of equilibria of DeGroot dynamics is $s_{\mathcal{G}}$-dimensional subspace of $\mathbb{R}^n$: if $s_{\mathcal{G}} > 1$, there are equilibria that are not consensus vectors.

$\mathcal{G}$ contains sinks (nodes i such that $W_{ij} = 0$ for all $j \neq i$), and at least one sink is reachable in $\mathcal{G}$ from every other node in $\mathcal{V}$: stubborn agents, their opinion remains constant (opinion leaders); other agents regular

$\mathcal{R}$ set of regular agents, $\mathcal{S}$ set of stubborn agents, $\mathcal{V} = \mathcal{R} \cup \mathcal{S}$
$Q \in \mathbb{R}^{\mathcal{R}\times\mathcal{R}}$ square block of P, $B \in \mathbb{R}^{\mathcal{R}\times\mathcal{S}}$ rectangular block of P, vector $y \in \mathbb{R}^{\mathcal{R}}$ opinions of regular agents, vector $u \in \mathbb{R}^{\mathcal{S}}$ opinions of stubborn agents (exogenous input)

$$y(t+1) = Qy(t) + Bu$$

Opinion dynamics with stubborn agents: theorem

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph with a non-empty set $\mathcal{S} \subseteq \mathcal{V}$ of sinks, at least one sink $s \in \mathcal{S}$ is reachable from every node $i \in \mathcal{V}$. For any initial opinion vector $x(0) \in \mathbb{R}^n$, the DeGroot opinion dynamics converges to an equilibrium opinion vector $x \in \mathbb{R}^n$ with

$$\begin{aligned} x_i &= \sum_{j \in \mathcal{V}} P_{ij} x_j \qquad i \in \mathcal{R}\,, \\ x_s &= x_s(0) \qquad s \in \mathcal{S}\,. \end{aligned}$$

Equivalently, the dynamics

$$y(t+1) = Qy(t) + Bu$$

of the regular agents' opinions vector $y(t) \in \mathbb{R}^{\mathcal{R}}$ converge to

$$y = (I - Q)^{-1} Bu = \sum_{k \geq 0} Q^k Bu$$

where $u \in \mathbb{R}^{\mathcal{S}}$ is the vector of the stubborn agents opinions.

Opinion dynamics with stubborn agents: disagreement 1

Discrete Laplace equation on $\mathcal{G}$ with boundary conditions on $\mathcal{S}$, or Dirichlet problem: multiplying by the out-degree,

$$\sum_{j \in \mathcal{V}} L_{ij} x_j = 0 \qquad i \in \mathcal{R}\,,$$
$$x_s = x_s(0) \qquad s \in \mathcal{S}\,.$$

If at least two stubborn nodes with different opinion, solution not constant over $\mathcal{R}$: no consensus is reached.

Disagreement among the regular agents persists and different regular agents reach a different equilibrium opinion, depending on their position in the network relative to the stubborn agents

Opinion dynamics with stubborn agents: disagreement 2

Equilibrium opinion x_i of a regular agent i is convex combination of opinions x_s of stubborn agents, weighted by a coefficient γ_s^i (normalized weight of all walks in $\mathcal{G}$ that start in the regular node i and terminate in the stubborn nodes s, without ever passing through any stubborn node in any intermediate step):

$$x_i = \sum_{s \in \mathcal{S}} \gamma_s^i x_s \,, \qquad \gamma_s^i = \sum_{k \geq 0} \sum_{\substack{i_1, \ldots, i_{k-1} \in \mathcal{R} \\ i_0 = i, i_k = s}} \prod_{1 \leq h \leq k} P_{i_{h-1} i_h} \,, \qquad i \in \mathcal{R}$$

The normalized weight of a link is smaller than 1: longer paths have smaller weight.

If a regular agent i is closer to one stubborn agent s, so that there are more shorter-length higher-weight paths connecting i to s than to any other stubborn node in $\mathcal{S}$, then the equilibrium opinion x_i tends to be biased towards x_s

Opinion dynamics with stubborn agents: example

Consider a network with $\mathcal{V} = \{1, \ldots, n\}$ and a line topology, with two stubborn agents placed in the two extremes.

$W \in \mathbb{R}^{n \times n}$ with $W_{1,1} = W_{n,n} = 1$, $W_{i,i+1} = W_{i,i-1} = 1$ for all $1 < i < n$, and $W_{ij} = 0$ for all other pair (i, j).

Stubborn nodes 1 and n have opinion $x_1 = 0$ and $x_n = 1$, respectively.

Then, the solution is

$$x_i = \frac{i-1}{n-1}, \qquad i = 1, 2, \ldots, n$$

The equilibrium opinions of the regular agents linearly interpolate between the extreme values of the stubborn agents.

Distributed algorithm to compute PageRank centrality 1

Parameter $\beta \in (0,1)$, intrinsic centrality μ, P normalized adjacency matrix of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$: PageRank centrality vector

$$\pi^{(\beta)} = \beta \left[I - (1-\beta) P^\top \right]^{-1} \mu$$

Add to $\mathcal{G}$ a sink node s with a self-loop of arbitrary positive weight, no outgoing links, incoming links of weight βw_i from every node $i \in \mathcal{V}$, with w_i out-degree: normalized adjacency matrix of the enlarged graph with node set $\mathcal{V} \cup \{s\}$ has $\mathcal{V} \times \mathcal{V}$ block $Q = (1-\beta)P$. Inflow vector $\lambda = \beta\mu$.

The linear system

$$y(t+1) = Q^\top y(t) + \lambda = (1-\beta) P^\top y(t) + \beta\mu$$

converges to

$$y = (I - Q^\top)^{-1} \lambda = \left[I - (1-\beta) P^\top \right]^{-1} \beta\mu = \pi^{(\beta)}$$

Distributed algorithm to compute PageRank centrality 2

Update rule

$$y_i(t+1) = (1-\beta)\sum_j P_{ji} y_j(t) + \beta \mu_i$$

iterative algorithm for the distributed computation of $\pi^{(\beta)}$
(each node needs information on in-neighbor states; to compute normalized weights $P_{ji} = W_{ji}/w_j$, also identity and out-degree of in-neighbors)

Linear compartmental systems: mass conservation

Nodes of connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ are cells/compartments that contain amount x_i of a commodity, which can flow from any cell i to out-neighbor cells of i

$x_i(t)$ in cell i varies in time due to unbalances between the sum of the inflows f_{ji} from other cells j and the sum of the outflows f_{ij} towards other cells j

External inflow λ_i into cell i (if present, open compartmental system; closed if $\lambda_i = 0$ for all i)

Dynamics of closed compartmental system (discrete time and continuous time), $i \in \mathcal{V}$

$$x_i(t+1) = x_i(t) + \sum_j f_{ji}(t) - \sum_j f_{ij}(t), \quad \dot{x}_i(t) = \sum_j f_{ji}(t) - \sum_j f_{ij}(t)$$

$$\text{Cell-to-cell flows} \qquad f_{ij}(t) = P_{ij} x_i(t), \qquad f_{ij}(t) = W_{ij} x_i(t),$$

$$x(t+1) = P^\top x(t), \qquad \dot{x} = -L^\top x$$

Linear compartmental systems (closed): theorem

Total mass in the system constant in time: $P\mathbb{1} = \mathbb{1}$ and $L\mathbb{1} = 0$,

$$\mathbb{1}^\top x(t+1) = \mathbb{1}^\top P^\top x(t) = \mathbb{1}^\top x(t)\,, \qquad \mathbb{1}^\top \dot{x}(t) = -\mathbb{1}^\top L^\top x(t) = 0\,.$$

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ connected graph, $w = W\mathbb{1}$ degree vector, $D = \operatorname{diag}(w)$, $P = D^{-1}W$ and $L = D - W$ normalized weight matrix and Laplacian matrix, $\pi = P^\top \pi$ unique invariant probability vector of $P^\top$ and $\bar{\pi} = D^{-1}\pi/(\mathbb{1}^\top D^{-1}\pi)$.

For every initial condition $x(0) \in \mathbb{R}^n_+$,

- if $\mathcal{G}$ is aperiodic, the discrete-time compartmental system $x(t+1) = P^\top x(t)$ converges to a limit

$$\lim_{t\to+\infty} x(t) = \mathbb{1}^\top x(0)\pi$$

- the continuous-time compartmental system $\dot{x} = -L^\top x$ converges to a limit

$$\lim_{t\to+\infty} x(t) = \mathbb{1}^\top x(0)\bar{\pi}$$

Linear compartmental systems (open)

Linear compartmental systems with a subset $\mathcal{S} \subseteq \mathcal{V}$ of sinks and constant (time-independent) inflows $\lambda_i > 0$ in nodes $i \in \mathcal{R} = \mathcal{V} \setminus \mathcal{S}$

$$x(t+1) = P^\top x(t) + \lambda \,, \qquad \dot{x} = -L^\top x + \lambda$$

Positive constant inflow in some node: total mass in the system blows up as time grows large.
Focus on a subnetwork!

Linear compartmental systems (open): theorem

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ graph containing a non-empty set $\mathcal{S} \subseteq \mathcal{V}$ of sinks; from every node $i \in \mathcal{R} = \mathcal{V} \setminus \mathcal{S}$ at least one sink $s \in \mathcal{S}$ is reachable. $y(t) \in \mathbb{R}^{\mathcal{R}}$ vector of the densities in the regular nodes, projection of $x(t)$ onto the set $\mathcal{R}$ (i.e., $y_i(t) = x_i(t)$ for all $i \in \mathcal{R}$); Q $\mathcal{R} \times \mathcal{R}$ block of P; M $\mathcal{R} \times \mathcal{R}$ block of L.

For every initial condition $y(0) \in \mathbb{R}^{\mathcal{R}}$,

- discrete-time compartmental system $y(t+1) = Q^\top y(t) + \lambda$ converges to

$$\lim_{t \to +\infty} y(t) = (I - Q^\top)^{-1} \lambda$$

- continuous-time compartmental system $\dot{y} = -M^\top y + \lambda$ converges to

$$\lim_{t \to +\infty} y(t) = (M^\top)^{-1} \lambda$$

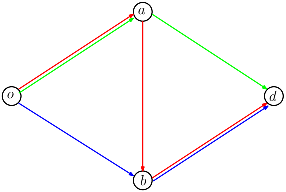
What we learned today

- Distributed averaging systems (e.g. sensor networks), both discrete- and continuous-time
- De Groot opinion dynamics: equilibria, convergence, consensus
- Distributed averaging over non-connected networks: stubborn nodes, from consensus to disagreement
- Google PageRank centrality
- Closed linear compartmental systems: mass conservation
- Open linear compartmental systems: inflows and outflows

SC42100 – Academic year 2019/2020

Networked and Distributed Control Systems

Connectivity and network flows



Giulia Giordano

University of Trento

(Visiting Prof. at TU Delft: g.giordano@tudelft.nl)

(Special thanks to **Giacomo Como**!)

From connectivity to Degree of Connectivity

Connectivity: graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ connected if there is a path from every node $i \in \mathcal{V}$ to any other node $j \in \mathcal{V}$

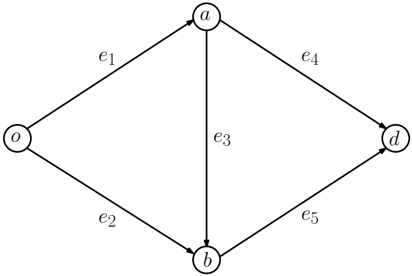
$\rightarrow$ binary property ($\mathcal{G}$ either connected or not)

Degree of connectivity?

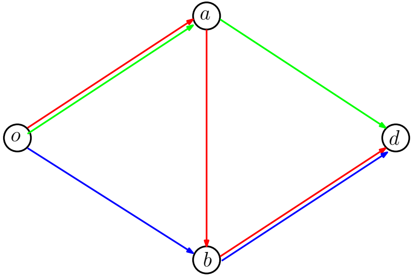
There might be multiple distinct o-d paths from node o (origin) to node d (destination):

$$p = (o = i_0, i_1, \ldots, i_{k-1}, i_k = d)$$

Example



Directed graph
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$
$\mathcal{V} = \{o, a, b, d\}$
$\mathcal{E} = \{e_1, e_2, e_3, e_4, e_5\}$



Three distinct *o*-*d* paths:
$p^{(1)} = (o, a, d)$
$p^{(2)} = (o, a, b, d)$
$p^{(3)} = (o, b, d)$

Link-path, node-link, node-path incidence matrices

- Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
- $\mathcal{P}$: set of all o-d paths in $\mathcal{G}$

$m = |\mathcal{P}|$ o-d paths, $l = |\mathcal{E}|$ (directed) links, $n = |\mathcal{V}|$ nodes

- Link-path incidence matrix $A \in \{0,1\}^{l \times m}$:

$$A_{ep} = \begin{cases} 1 & \text{if} \quad \text{link } e \text{ is along path } p \\ 0 & \text{if} \quad \text{link } e \text{ is not along path } p\,. \end{cases}$$

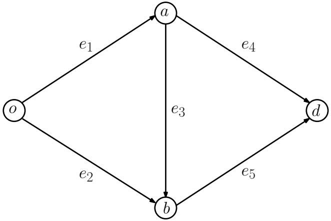
- Node-link incidence matrix $B \in \{-1, 0, +1\}^{n \times l}$:

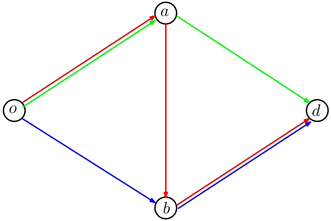
$$e = (i, j)\,,\ k \neq i, j \quad \Longrightarrow \quad B_{ie} = +1\,, \quad B_{je} = -1\,, \quad B_{ke} = 0$$

- Node-path incidence matrix: multiply B by A, $(BA) \in \{-1, 0, +1\}^{n \times m}$. For every node i and o-d path p:

$$(BA)_{ip} = \begin{cases} +1 & \text{if} \quad i = o \\ -1 & \text{if} \quad i = d \\ \ \ 0 & \text{if} \quad i \neq o, d \end{cases}$$

Example: link-path, node-link, node-path incidence matrices





$$A = \begin{array}{c@{}c} \begin{array}{ccc} \color{green}{p^{(1)}} & \color{red}{p^{(2)}} & \color{blue}{p^{(3)}} \end{array} & \\ \left[\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right] & \begin{array}{c} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{array} \end{array} \qquad B = \begin{array}{c@{}c} \begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \end{array} & \\ \left[\begin{array}{ccccc} +1 & +1 & 0 & 0 & 0 \\ -1 & 0 & +1 & +1 & 0 \\ 0 & -1 & -1 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right] & \begin{array}{c} o \\ a \\ b \\ d \end{array} \end{array}$$

$$BA = \begin{array}{c@{}c} \begin{array}{ccc} \color{green}{p^{(1)}} & \color{red}{p^{(2)}} & \color{blue}{p^{(3)}} \end{array} & \\ \left[\begin{array}{ccc} +1 & +1 & +1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{array}\right] & \begin{array}{c} o \\ a \\ b \\ d \end{array} \end{array}$$

Node-connectivity and link-connectivity

Different o-d paths may share intermediate nodes or links

$$p^{(1)} = (o = i_0^{(1)}, i_1^{(1)}, \ldots, i_{l_1}^{(1)} = d)$$
$$p^{(2)} = (o = i_0^{(2)}, i_1^{(2)}, \ldots, i_{l_2}^{(2)} = d)$$

- node-independent if they share no *intermediate* node ($i_h^{(1)} \neq i_k^{(2)}$ for all $1 \leq h < l_1$ and $1 \leq k < l_2$)
- link-independent if they share no link ($(i_{h-1}^{(1)}, i_h^{(1)}) \neq (i_{k-1}^{(2)}, i_k^{(2)})$ for all $1 \leq h \leq l_1$ and $1 \leq k \leq l_2$)

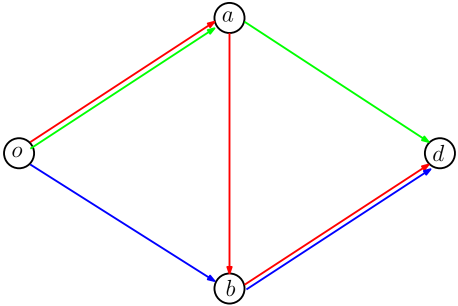
Node-connectivity $c_{\text{node}}(o, d)$: $\#$ node-independent o-d paths

Link-connectivity: $c_{\text{link}}(o, d)$ $\#$ link-independent o-d paths

For an unweighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$,

$$c_{\text{node}}(\mathcal{G}) = \min_{o \neq d \in \mathcal{V}} c_{\text{node}}(o, d), \quad c_{\text{link}}(\mathcal{G}) = \min_{o \neq d \in \mathcal{V}} c_{\text{link}}(o, d)$$

Example: node-connectivity and link-connectivity



$p^{(1)}$ and $p^{(3)}$ are both node- and link-independent

$p^{(2)}$ is neither node- nor link-independent from either $p^{(1)}$ or $p^{(3)}$

$c_{\text{node}}(o, d) = c_{\text{link}}(o, d) = 2$

$c_{\text{node}}(\mathcal{G}) = c_{\text{link}}(\mathcal{G}) = 0$ because $\mathcal{G}$ is not connected

Menger's Theorem

How many nodes and links must we remove from a graph to disconnect two nodes?

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ unweighted graph, $i \neq j$ two distinct nodes.

- Minimum number of nodes that has to be removed from $\mathcal{G}$ in order for j not to be reachable from i is $c_{\text{node}}(i, j)$.
- Minimum number of links that has to be removed from $\mathcal{G}$ in order for j not to be reachable from i is $c_{\text{link}}(i, j)$.
- Minimum number of nodes that have to be removed from $\mathcal{G}$ in oder to disconnect it is $c_{\text{node}}(\mathcal{G})$.
- Minimum number of links that have to be removed from $\mathcal{G}$ in oder to disconnect it is $c_{\text{link}}(\mathcal{G})$.

Special case of a more general result: max-flow min-cut theorem.

Network flows

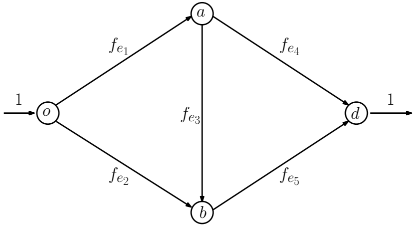
- graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, node-link incidence matrix $B \in \{-1, 0, +1\}^{\mathcal{V}\times\mathcal{E}}$

$$e = (i, j)\,,\ k \neq i, j \quad \Longrightarrow \quad B_{ie} = +1\,, \quad B_{je} = -1\,, \quad B_{ke} = 0$$

- exogenous net flow vector $\nu \in \mathbb{R}^{\mathcal{V}}$ such that $\sum_i \nu_i = 0$
- network flow is a vector $f \in \mathbb{R}^{\mathcal{E}}$ such that

$$f \geq 0 \qquad Bf = \nu$$

- typically several feasible solutions when problem feasible
- Example



$$B = \begin{array}{c} \begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \end{array} \\ \begin{bmatrix} +1 & +1 & 0 & 0 & 0 \\ -1 & 0 & +1 & +1 & 0 \\ 0 & -1 & -1 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} \end{array}$$

Network flows - inflows and outflows

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ with no self-loops

λ, $\mu \in \mathbb{R}^n_+$ inflows and outflows in the nodes (nonnegative vectors)

- Mass conservation

$$\mathbb{1}^\top \lambda = \mathbb{1}^\top \mu = \tau$$

total inflow $\mathbb{1}^\top \lambda = \sum_{i \in \mathcal{V}} \lambda_i$ equals total outflow $\mathbb{1}^\top \mu = \sum_{i \in \mathcal{V}} \mu_i$: value τ throughput

- Balance constraint

$$\lambda^\top \mu = 0$$

$\lambda_i, \mu_i \geq 0 \; \forall i \in \mathcal{V} \Longrightarrow$ at least one (possibly both) of λ_i and μ_i is 0

If both $\lambda_i, \mu_i > 0$, subtract $\min\{\lambda_i, \mu_i\}$ from both: equivalent net flow $\lambda_i - \mu_i$ in node i satisfies the constraint.

Nodes i such that $\lambda_i > 0$ and $\mu_i = 0$: *sources*, *origins*, *generators*

Nodes i such that $\lambda_i = 0$ and $\mu_i > 0$: *sinks*, *destinations*, *loads*

Network flows - flow vectors

Flow vectors: nonnegative vectors $f \in \mathbb{R}_+^{\mathcal{E}}$, whose entries $f_{(i,j)}$ are the flows along the links $(i,j) \in \mathcal{E}$

- The total inflow in a node i (external inflow λ_i plus flows $f_{(j,i)}$ from in-neighbors j) must equal the total outflow from i (external outflow μ_i plus flows $f_{(i,j)}$ towards out-neighbors j):

$$\lambda_i + \sum_{j:(j,i)\in\mathcal{E}} f_{(j,i)} = \sum_{j:(i,j)\in\mathcal{E}} f_{(i,j)} + \mu_i\,, \qquad i \in \mathcal{V}$$

- Compact form using node-link incidence matrix:

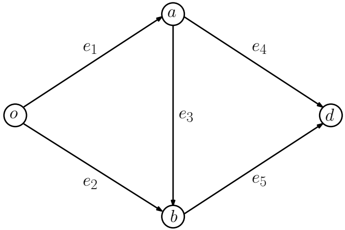
$$Bf = \lambda - \mu$$

- Flows from a single origin o to a single destination d: *o-d* flows, with external inflow vector $\lambda = \delta^{(o)}$ and outflow vector $\mu = \delta^{(d)}$

$\delta^{(i)}$: a 1 entry in the i-th position and all other entries equal to 0

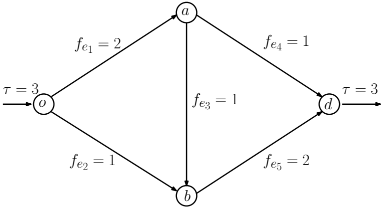
Example: network flows

o-*d* flow: nonnegative vector $f = (f_{e_1}, f_{e_2}, f_{e_3}, f_{e_4}, f_{e_5})$
satisfying the flow balance:

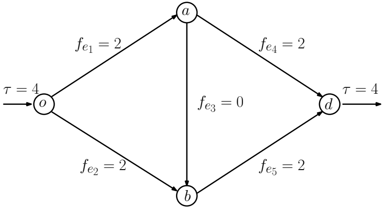


$$\tau = f_{e_1} + f_{e_2}$$
$$f_{e_1} = f_{e_3} + f_{e_4}$$
$$f_{e_2} + f_{e_3} = f_{e_5}$$
$$f_{e_4} + f_{e_5} = \tau$$

Two possible *o*-*d* flows:



lower throughput $\tau = 3$



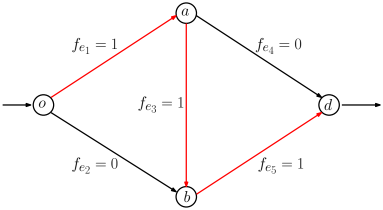
higher throughput $\tau = 4$

Network flows - unitary o-d flows

$\forall$ o-d paths $p \in \mathcal{P}$, $A\delta^{(p)} \in \mathbb{R}^{|\mathcal{E}|}$:

- p-th column of the link-path incidence matrix
- has entries 1 for links along the path p and 0 otherwise
- is a unitary o-d flow (a flow from o to d of throughput 1):

$$BA\delta^{(p)} = \delta^{(o)} - \delta^{(d)}$$



$p^{(2)}$-th column of the link-path incidence matrix A for the graph is a unitary o-d flow

Network flows - flow assignment

Nonnegative vector $z \in \mathbb{R}^{\mathcal{P}}$, $z_p \geq 0$ aggregate flow on *o*-*d* path p.

$$f = \sum_{p \in \mathcal{P}} z_p A \delta^{(p)} = Az$$

is an *o*-*d* flow of throughput τ:

$$f \geq 0\,, \qquad Bf = BAz = \tau(\delta^{(o)} - \delta^{(d)})\,, \qquad \tau = \mathbb{1}^\top z$$

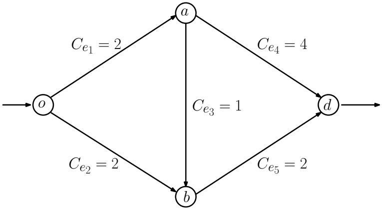
Assign flows z to *o*-*d* paths (and cycles) in the graph
$\rightarrow$ unique *o*-*d* flow $f = Az$ on the links
(useful to construct feasible flows)

Given network flow f on the links
$\rightarrow$ there is a possibly (and typically) non-unique assignment of flows to both *o*-*d* paths and cycles in the graph that induces f
(Flow Decomposition Theorem)

Capacity

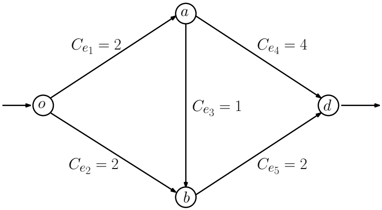
Link $e \in \mathcal{E}$ has capacity $C_e > 0$:
maximum flow allowable through the link.

Vector of all link capacities: $C \in \mathbb{R}^{\mathcal{E}}$.



Maximum throughput τ from node o to node d that can be achieved by a flow f without violating the link capacity constraints?

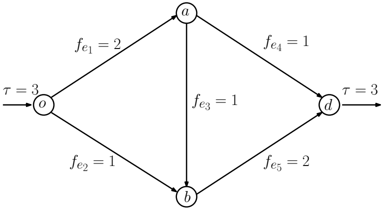
Example: maximum throughput with capacity constraints

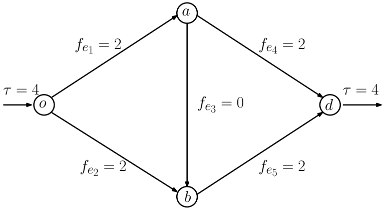


Maximize τ over link flows $f_{e_1}, f_{e_2}, f_{e_3}, f_{e_4}, f_{e_5}$ and throughput τ s.t.

$$\tau = f_{e_1} + f_{e_2}, \quad f_{e_1} = f_{e_3} + f_{e_4}, \quad f_{e_2} + f_{e_3} = f_{e_5}, \quad f_{e_4} + f_{e_5} = \tau,$$

$$\tau \geq 0\,, \qquad 0 \leq f_{e_k} \leq C_{e_k}\,, \qquad 1 \leq k \leq 5$$





Max flow problem

Directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.

- Link capacity vector $C \in \mathbb{R}^{\mathcal{E}}$, $C > 0$
- Link flows vector $f \in \mathbb{R}_+^{\mathcal{E}}$
- Throughput $\tau \geq 0$, total flow through the network from node o to node d, associated with f

Consider two distinct nodes o and d. Maximum flow problem:

$$
\begin{array}{rcll}
\tau^*_{o,d} & = & \max \tau & \\
\text{s.t.} & & \tau \geq 0 & \text{throughput nonnegativity} \\
 & & 0 \leq f \leq C & \text{nonnegativity and capacity constraints} \\
 & & Bf = \tau(\delta^{(o)} - \delta^{(d)}) & \text{mass conservation}
\end{array}
$$

Linear program: objective function and constraints are linear functions of the variables

Flow satisfying the constraints: feasible flow. Set of feasible flows nonempty: it always contains flow $f = 0$ with throughput $\tau = 0$

Min cut capacity

o-d cut: partition of the node set $\mathcal{V}$ in two subsets, $\mathcal{U}$ and $\mathcal{V} \setminus \mathcal{U}$, with $o \in \mathcal{U}$ and $d \in \mathcal{V} \setminus \mathcal{U}$

Capacity of an *o-d* cut $\mathcal{U}$ is the aggregate capacity of the links from $\mathcal{U}$ to $\mathcal{V} \setminus \mathcal{U}$:

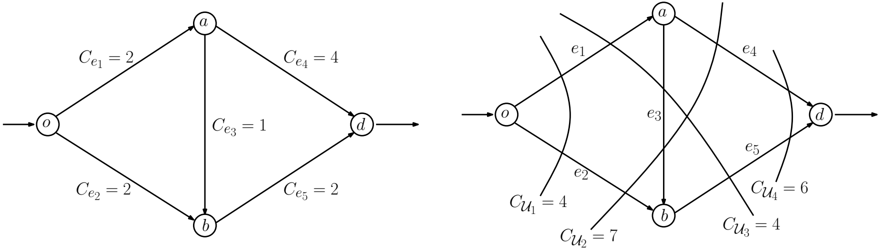
$$C_{\mathcal{U}} := \sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{V} \setminus \mathcal{U}} C_{ij}$$

Min-cut capacity: minimum capacity among all *o-d* cuts

$$C^*_{o,d} = \min_{\substack{\mathcal{U} \subseteq \mathcal{V} \\ o \in \mathcal{U},\, d \notin \mathcal{U}}} C_{\mathcal{U}}$$

Bottleneck: cut $\mathcal{U}$ with minimal capacity, $C_{\mathcal{U}} = C^*_{o,d}$

Example: cut capacity



Four o-d cuts	with capacities
$\mathcal{U}_1 = \{o\}$,	$C_{\mathcal{U}_1} = C_{e_1} + C_{e_2} = 4$,
$\mathcal{U}_2 = \{o, a\}$,	$C_{\mathcal{U}_2} = C_{e_2} + C_{e_3} + C_{e_4} = 7$,
$\mathcal{U}_3 = \{o, b\}$,	$C_{\mathcal{U}_3} = C_{e_1} + C_{e_5} = 4$,
$\mathcal{U}_4 = \{o, a, b\}$	$C_{\mathcal{U}_4} = C_{e_4} + C_{e_5} = 6$

Min-cut capacity: $C^*_{o,d} = 4$
Bottlenecks (minimal capacity cuts): $\mathcal{U}_1$ and $\mathcal{U}_3$

Max-flow min-cut theorem

How do we guarantee that a flow vector achieves the maximum throughput $\tau^*_{o,d}$ from an origin node o to a destination node d?

Relate $\tau^*_{o,d}$ to geometrical properties of the graph $\mathcal{G}$

Max-flow min-cut theorem: maximum throughput $\tau^*_{o,d}$ from o to d (solution of the linear program) coincides with the minimum cut capacity $C^*_{o,d}$ among all o-d cuts:

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ directed graph, C vector of positive link capacities. For every two nodes $o \neq d \in \mathcal{V}$,

$$\tau^*_{o,d} = C^*_{o,d}$$

If the link capacities are all integer-valued, then a feasible maximum throughput flow can be constructed such that the flow on every link is an integer value.

Max-flow min-cut theorem and Menger's theorem

Max-flow min-cut: the minimum total capacity to be removed from the network so as to make node d not reachable from node o coincides with the min-cut capacity $C^*_{o,d}$

If the link capacity values are 0 or 1 (either remove or keep links), then integer-valued feasible flows satisfy $f_{(i,j)} = 0, 1$ for every link $(i,j) \in \mathcal{E}$. Set of links (i,j) such that $f_{i,j} = 1$ is the union of link-disjoint paths from o to d:

$\rightarrow$ max-flow min-cut theorem reduces to Menger's theorem

Max-flow min-cut theorem: proof $\tau^*_{o,d} = C^*_{o,d}$

Two steps

- Show that $\tau^*_{o,d} \leq C^*_{o,d}$: no feasible flow can have throughput larger than the min-cut capacity (easy)
- Show that $\tau^*_{o,d} \geq C^*_{o,d}$: there exists a feasible flow with throughput equal to the min-cut capacity (hard)

Max-flow min-cut theorem: proof $\tau^*_{o,d} \leq C^*_{o,d}$ (1)

$\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$, $\quad \mathcal{E}_{\mathcal{A}\to\mathcal{B}} = \{(i,j) \in \mathcal{E} : \; i \in \mathcal{A}\,,\; j \in \mathcal{B}\}$ set of links from some node in $\mathcal{A}$ to some node in $\mathcal{B}$

$\forall$ feasible flows f from o to d with throughput τ and $\forall$ o-d cuts $\mathcal{U}$,

$$\tau + \sum_{e \in \mathcal{E}_{\mathcal{V}\setminus\mathcal{U}\to\mathcal{U}}} f_e = \sum_{e \in \mathcal{E}_{\mathcal{U}\to\mathcal{V}\setminus\mathcal{U}}} f_e$$

Mass conservation: outflow from $\mathcal{U}$ equals the sum of inflow τ in the origin o plus flow from the rest of the network $\mathcal{V}\setminus\mathcal{U}$ towards $\mathcal{U}$

Indeed, by summing for all $i \in \mathcal{U}$ node-wise mass conservation laws

$$\lambda_i + \sum_{e=(j,i)\in\mathcal{E}} f_e = \sum_{e=(i,j)\in\mathcal{E}} f_e + \mu_i$$

$$\text{we get } \sum_{i\in\mathcal{U}} \lambda_i + \sum_{e\in\mathcal{E}_{\mathcal{V}\to\mathcal{U}}} f_e = \sum_{e\in\mathcal{E}_{\mathcal{U}\to\mathcal{V}}} f_e + \sum_{i\in\mathcal{U}} \mu_i$$

Since $\mathcal{U}$ is an o-d cut, $\sum_{i\in\mathcal{U}} \lambda_i = \tau$ and $\sum_{i\in\mathcal{U}} \mu_i = 0$

$$\tau + \sum_{e\in\mathcal{E}_{\mathcal{V}\to\mathcal{U}}} f_e - \sum_{e\in\mathcal{E}_{\mathcal{U}\to\mathcal{U}}} f_e = \sum_{e\in\mathcal{E}_{\mathcal{U}\to\mathcal{V}}} f_e - \sum_{e\in\mathcal{E}_{\mathcal{U}\to\mathcal{U}}} f_e$$

Max-flow min-cut theorem: proof $\tau^*_{o,d} \leq C^*_{o,d}$ (2)

We have seen that

$$\tau + \sum_{e \in \mathcal{E}_{\mathcal{V}\setminus\mathcal{U} \to \mathcal{U}}} f_e = \sum_{e \in \mathcal{E}_{\mathcal{U} \to \mathcal{V}\setminus\mathcal{U}}} f_e$$

Since $0 \leq f_e \leq C_e$ for the flow on every link e,

$$\sum_{e \in \mathcal{E}_{\mathcal{U} \to \mathcal{V}\setminus\mathcal{U}}} C_e \geq \sum_{e \in \mathcal{E}_{\mathcal{U} \to \mathcal{V}\setminus\mathcal{U}}} f_e = \tau + \sum_{e \in \mathcal{E}_{\mathcal{V}\setminus\mathcal{U} \to \mathcal{U}}} f_e \geq \tau$$

If we choose the cut $\mathcal{U}$ so that it is a bottleneck,
$\sum_{e \in \mathcal{E}_{\mathcal{U} \to \mathcal{V}\setminus\mathcal{U}}} C_e = C^*_{o,d}$ is the min-cut capacity and $C^*_{o,d} \geq \tau^*_{o,d}$.

Max-flow min-cut theorem: proof $\tau^*_{o,d} \geq C^*_{o,d}$

We need to construct a feasible flow f from o to d with throughput τ equal to the min-cut capacity $C^*_{o,d}$.

An iterative algorithm due to Ford and Fulkerson does this in a finite number of steps, by starting with a trivial flow $f^{(0)} = 0$ with throughput $\tau^{(0)} = 0$ and capacity vector $C^{(0)} = C$ and then constructing a feasible flow for which $\tau^*_{o,d} = C^*_{o,d}$.

(details on the algorithm are in the lecture notes)

If the link capacities are all positive integers, then the flow vector constructed by the algorithm has integer entries.

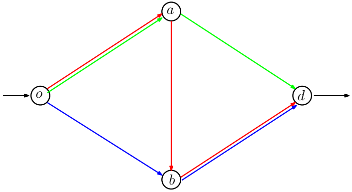
What we learned today

- Degree of connectivity
- Incidence matrix and properties
- Network flows, capacity, throughput
- Flows and cuts: Menger's Theorem and Max-flow Min-cut Theorem

SC42100 – Academic year 2019/2020

Networked and Distributed Control Systems

Network Flow Optimization



Giulia Giordano

University of Trento

(Visiting Prof. at TU Delft: `g.giordano@tudelft.nl`)

(Special thanks to **Giacomo Como** and **Gustav Nilsson**!)

Network flows

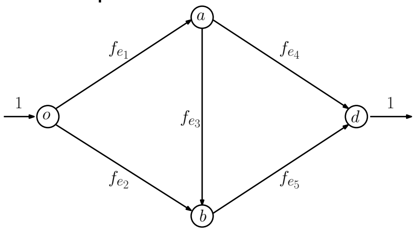
- graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, node-link incidence matrix $B \in \{-1, 0, +1\}^{\mathcal{V}\times\mathcal{E}}$

$$e = (i, j)\,,\ k \neq i, j \quad \Longrightarrow \quad B_{ie} = +1\,, \quad B_{je} = -1\,, \quad B_{ke} = 0$$

- exogenous net flow vector $\nu \in \mathbb{R}^{\mathcal{V}}$ such that $\sum_i \nu_i = 0$
- network flow is a vector $f \in \mathbb{R}^{\mathcal{E}}$ such that

$$f \geq 0 \qquad Bf = \nu$$

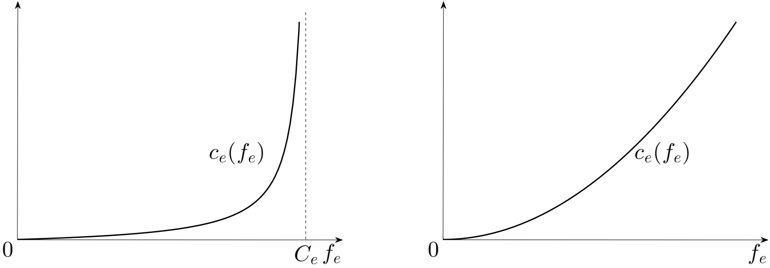
- typically several feasible solutions when problem feasible
- Example



$$B = \begin{array}{c} \begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \end{array} \\ \begin{bmatrix} +1 & +1 & 0 & 0 & 0 \\ -1 & 0 & +1 & +1 & 0 \\ 0 & -1 & -1 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} \end{array}$$

Network flow optimization

- convex nondecreasing cost functions $c_e(f_e)$ on every link $e \in \mathcal{E}$



- Either finite or infinite link flow capacity

$$C_e = \inf\{f_e \geq 0 : \, c_e(f_e) = +\infty\}$$

- network flow optimization problem

$$M(\nu) := \min_{\substack{f \geq 0 \\ Bf = \nu}} \sum_{e \in \mathcal{E}} c_e(f_e)$$

Example 1: Shortest path and optimal transport

$$c_e(f_e) = l_e f_e \qquad e \in \mathcal{E}$$

$$l_e = \text{ length of link } e$$

- if $\nu = \delta^{(o)} - \delta^{(d)}$, shortest path
- if $\nu = \lambda - \mu$ with $\lambda, \mu \in \mathbb{R}_+^{\mathcal{V}}$, optimal transport

Example 2: Power dissipation

Undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, link resistances $R_{ij} = R_{ji} > 0$

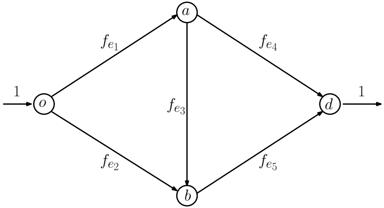
$$c_{(i,j)}(f_{(i,j)}) = \frac{R_{ij}^{\alpha}}{\alpha + 1} |f_{(i,j)}|^{\alpha+1} \qquad (i,j) \in \mathcal{E}$$

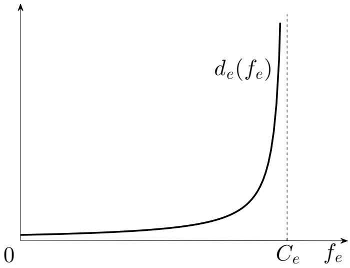
- Power dissipation:

$$\sum_{(i,j)\in\mathcal{E}} \frac{R_{ij}^{\alpha}}{\alpha + 1} |f_{(i,j)}|^{\alpha+1}$$

- $\alpha = 1 \quad \Rightarrow \quad$ direct current (DC) power networks
- $\alpha = 2 \quad \Rightarrow \quad$ gas networks
- $\alpha = 1/1.85 \simeq 0.54 \quad \Rightarrow \quad$ water networks

Example 3: System optimum traffic assignment (SO-TAP)





- convex nondecreasing delay functions $d_e(f_e)$. E.g.,

$$d_e(f_e) = \frac{l_e}{1 - f_e/C_e}\,, \qquad e \in \mathcal{E}$$

- corresponding cost

$$c_e(f_e) = f_e \cdot d_e(f_e)$$

is convex nondecreasing

Lagrangian techniques and duality

- network flow optimization problem

$$M(\nu) := \min_{\substack{f \geq 0 \\ Bf = \nu}} \sum_{e \in \mathcal{E}} c_e(f_e)$$

- Lagrange multipliers γ_i for every node $i \in \mathcal{V}$
- Lagrangian function

$$\begin{aligned} L(f, \gamma, \nu) &= \sum_{(i,j) \in \mathcal{E}} c_{(i,j)}(f_{(i,j)}) + \sum_{i \in \mathcal{V}} \gamma_i \left(\sum_{j:(j,i) \in \mathcal{E}} f_{(j,i)} - \sum_{j:(i,j) \in \mathcal{E}} f_{(i,j)} + \nu_i \right) \\ &= \sum_{(i,j) \in \mathcal{E}} \left(c_{(i,j)}(f_{(i,j)}) - f_{(i,j)}(\gamma_i - \gamma_j) \right) + \sum_{i \in \mathcal{V}} \gamma_i \cdot \nu_i \end{aligned}$$

- dual optimization

$$D(\gamma, \nu) := \min_{f \geq 0} L(f, \gamma, \nu)$$

Lagrangian techniques and duality

- KKT conditions (necessary and sufficent)

$$c'_e(f^*_{(i,j)}) \begin{cases} = \gamma_i - \gamma_j & \text{if} \quad f^*_{(i,j)} > 0 \\ \geq \gamma_i - \gamma_j & \text{if} \quad f^*_{(i,j)} = 0 \end{cases} , \qquad (i,j) \in \mathcal{E} \,.$$

- to find the multipliers γ impose conservation of mass or solve dual optimization

$$\gamma^* \in \operatorname*{argmax}_{\gamma \in \mathbb{R}^{\mathcal{V}}} D(\gamma, \nu) \,,$$

- note that

$$D(\gamma, \nu) = \sum_{(i,j)} c^*_{(i,j)}(\gamma_i - \gamma_j) - \sum_i \gamma_i \cdot \nu_i$$

where the dual costs

$$c^*_e(y_e) = \sup_{f_e \geq 0} \{y_e f_e - c_e(f_e)\} \qquad e \in \mathcal{E}$$

represent the maximum profit that a link operator can make if it charges y_e per unit of flow and pays $c_e(f_e)$ to transport f_e units

Lagrangian techniques and duality

Theorem: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, nonincreasing costs $c_e(f_e)$, capacities C_e

(i) the problem is feasible, i.e., $M(\nu) < +\infty$, if and only if

$$\sum_{i \in \mathcal{U}} \nu_i < \sum_{\substack{(i,j) \in \mathcal{E}: \\ i \in \mathcal{U}\,,\ j \in \mathcal{V} \setminus \mathcal{U}}} C_{(i,j)}\,, \tag{1}$$

for every nonempty proper $\mathcal{U} \subset \mathcal{V}$.

If $c_e(f_e)$ convex and differentiable on $[0, +\infty)$, then for all ν s.t. (1),

(ii)

$$M(\nu) = \max_{\gamma \in \mathbb{R}^{\mathcal{V}}} D(\gamma, \nu)$$

flow vector f^* is optimal if and only if it satisfies the KKT conditions for Lagrange multipliers $\gamma = \gamma^*$;

(iii) if the optimal cost $M(\nu)$ is differentiable in ν, then

$$\frac{\partial}{\partial \nu_i} M(\nu) - \frac{\partial}{\partial \nu_j} M(\nu) = \gamma_i^* - \gamma_j^*$$

Example – KKT conditions for shortest path

$$c_e(f_e) = l_e f_e \qquad e \in \mathcal{E}$$

$$l_e = \text{ length of link } e$$

- $\nu = \delta^{(o)} - \delta^{(d)}$
- KKT conditions yield

$$l_{(i,j)} \begin{cases} = \gamma_i - \gamma_j & \text{if } f^*_{(i,j)} > 0 \\ \geq \gamma_i - \gamma_j & \text{if } f^*_{(i,j)} = 0 \end{cases} \qquad (i,j) \in \mathcal{E}$$

- For every $o - d$ path, $o = i_0, i_2, i_2, \dots, i_{k-1}, i_k = d$

$$\sum_{j=1}^{k} l_{(i_{j-1}, i_j)} \begin{cases} = \gamma_o - \gamma_d & \text{if } f^*_{(i_{j-1}, i_j)} > 0 & \text{for all } 1 \leq j \leq k \\ \geq \gamma_o - \gamma_d & \text{if } f^*_{(i_{j-1}, i_j)} = 0 & \text{for some } 1 \leq j \leq k \end{cases}$$

Example – KKT conditions for power dissipation

$$c_e(f_e) = \frac{R_e^\alpha}{\alpha+1}|f_e|^{\alpha+1} \qquad e \in \mathcal{E}$$

- For $f_e \geq 0$, $c_e'(f_e) = R_e f_e^\alpha$
- KKT conditions yield

$$f^*_{(i,j)} = \begin{cases} \frac{(\gamma_i - \gamma_j)^{1/\alpha}}{R_{ij}} & \text{if } \gamma_i > \gamma_j \\ 0 & \text{if } \gamma_i \leq \gamma_j \end{cases} \qquad (i,j) \in \mathcal{E}$$

- Netflow z_{ij} from node i to node j

$$z_{ij} = f^*_{(i,j)} - f^*_{(j,i)} = R_{ij}^{-1}|\gamma_i - \gamma_j|^{1/\alpha}\mathrm{sgn}(\gamma_i - \gamma_j)$$

- DC power networks ($\alpha = 1$)

$$z_{ij} = \frac{\gamma_i - \gamma_j}{R_{ij}}$$

Coincides with Ohm's law

User optimal traffic assignment

- Drivers, total amount τ, can choose different paths between o and d
- $\Gamma_{o,d}$ the set of all o-d paths
- Link-path incidence matrix $A^{(o,d)} \in \{0,1\}^{\mathcal{E}\times\Gamma_{o,d}}$

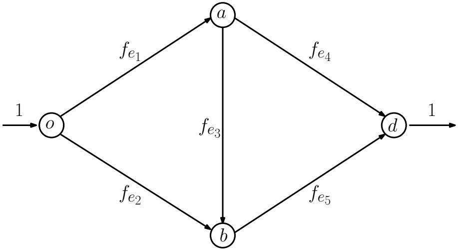
$$A^{(o,d)}_{ep} = \begin{cases} 1 & \text{if link } e \text{ is along path } p \\ 0 & \text{if link } e \text{ is not along path } p \end{cases}$$

- Path flow $z \in \mathbb{R}^{\Gamma_{o,d}}$, $\mathbb{1}'z = \tau$, $z \geq 0$
- Observe that

$$(BA^{(o,d)})_{ip} = \begin{cases} +1 & \text{if } i = o \\ -1 & \text{if } i = d \\ 0 & \text{if } i \neq o, d \end{cases}$$

- Link flow $f = A^{(o,d)}z$

Example – Link-path incidence matrix



$$A^{(o,d)} = \begin{array}{c c} \begin{array}{ccc} p^{(1)} & p^{(2)} & p^{(3)} \end{array} & \\ \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} & \begin{array}{c} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{array} \end{array} \quad B = \begin{array}{c c} \begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \end{array} & \\ \begin{bmatrix} +1 & +1 & 0 & 0 & 0 \\ -1 & 0 & +1 & +1 & 0 \\ 0 & -1 & -1 & 0 & +1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} & \begin{array}{c} o \\ a \\ b \\ d \end{array} \end{array}$$

Wardrop equilibrium

- Wardrop equilibrium $f^{(0)}$: The flow vector

$$f^{(0)} = A^{(o,d)} z$$

where $z \in \mathbb{R}^{\Gamma_{o,d}}$ is such that $z \geq 0$, $\mathbb{1}' z = \tau$, and for $p \in \Gamma_{o,d}$

$$z_p > 0 \quad \Longrightarrow \quad \underbrace{\sum_{e \in \mathcal{E}} A^{(o,d)}_{ep} d_e(f^{(0)}_e)}_{\substack{\text{total delay} \\ \text{on path } p}} \quad \leq \quad \underbrace{\sum_{e \in \mathcal{E}} A^{(o,d)}_{eq} d_e(f^{(0)}_e)}_{\substack{\text{total delay} \\ \text{on path } q}} \qquad \forall q \in \Gamma_{o,d}$$

- Interpretation: Each driver chooses their fastest path
- The Wardrop equilibrium can be computed as

$$M(\nu) := \min_{\substack{f \geq 0 \\ Bf = \tau(\delta^{(o)} - \delta^{(d)})}} \sum_{e \in \mathcal{E}} \int_0^{f_e} d_e(s) \mathrm{d}s$$

Price of Anarchy

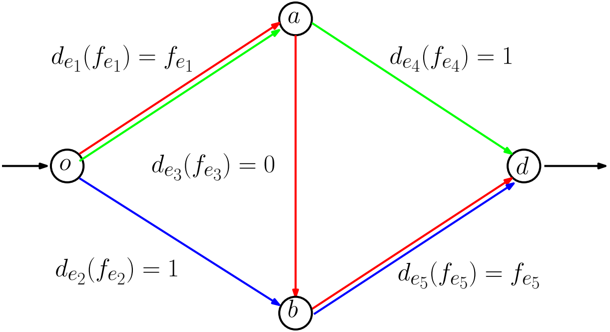
- *price of anarchy* associated to Wardrop equilibrium $f^{(0)}$ is

$$\mathrm{PoA}(0) = \frac{\sum_{e\in\mathcal{E}} f_e^{(0)} d_e(f_e^{(0)})}{\min_{\substack{f\geq 0 \\ Bf=\tau(\delta^{(o)}-\delta^{(d)})}} \sum_{e\in\mathcal{E}} f_e d_e(f_e)},$$

total delay at the Wardrop equilibrium / total delay at system optimum

- Observe: PoA ≥ 1

Example - Braess paradox



- Three paths $z_{p^{(1)}}$, $z_{p^{(2)}}$, $z_{p^{(3)}}$
- Wardrop equilibrium: $z_{p^{(1)}} = z_{p^{(2)}} = 0$, $z_{p^{(3)}} = 1$
- Social optimum: $z_{p^{(1)}} = z_{p^{(2)}} = 1/2$, $z_{p^{(3)}} = 0$

Toll design

- Let $\omega = (\omega_e)_{e\in\mathcal{E}}$ be vector of tolls
- The cost for each link $\omega_e + d_e(f_e)$
- Wardrop equilibrium with tolls

$$f^{(\omega)} = A^{(o,d)}z$$

where $z \in \mathbb{R}^{\Gamma_{o,d}}$ is such that $z \geq 0$, $\mathbb{1}'z = \tau$, and for $p \in \Gamma_{o,d}$

$$z_p > 0 \Longrightarrow \underbrace{\sum_{e\in\mathcal{E}} A^{(o,d)}_{ep}\left(d_e(f^{(\omega)}_e) + \omega_e\right)}_{\substack{\text{total delay}\\ \text{on path } p}} \leq \underbrace{\sum_{e\in\mathcal{E}} A^{(o,d)}_{eq}\left(d_e(f^{(\omega)}_e) + \omega_e\right)}_{\substack{\text{total delay}\\ \text{on path } q}}$$

- Can we find ω s.t. PoA(ω)$= 1$?

Toll design

Let $d_e(f_e)$ be strictly increasing, $f_e d_e(f_e)$ strictly convex,

$$f^{(\omega^*)} = \underset{Bf=\tau(\delta^{(o)}-\delta^{(d)}),\, f\geq 0}{\operatorname{argmin}} \sum_{e\in\mathcal{E}} \left(\int_0^{f_e} d_e(s) + \omega_e f_e \right)$$

and let f^* denote the social optimum, i.e.,

$$f^* = \underset{Bf=\tau(\delta^{(o)}-\delta^{(d)}),\, f\geq 0}{\operatorname{argmin}} \sum_{e\in\mathcal{E}} f_e d_e(f_e)\,.$$

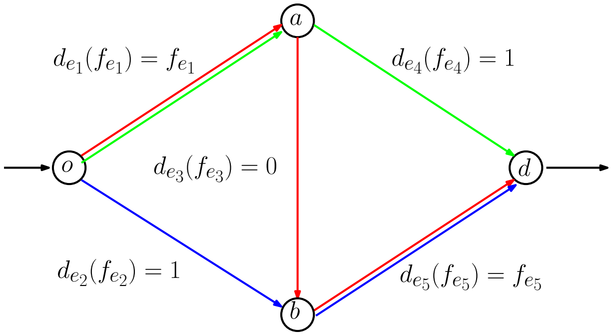
If we let

$$\omega_e = f_e^* d_e'(f_e^*)$$

then

$$f^{(\omega^*)} = f^*\,.$$

Example – Toll design



- Tolls $\omega_{e_1} = \omega_{e_5} = 1/2$, $\omega_{e_2} = \omega_{e_3} = \omega_{e_4} = 0$
- Wardrop equilibrium with tolls: $z_{p^{(1)}} = z_{p^{(2)}} = 1/2$, $z_{p^{(3)}} = 0$
- Social optimum: $z_{p^{(1)}} = z_{p^{(2)}} = 1/2$, $z_{p^{(3)}} = 0$

What we learned today

- Network flow optimization: shortest path, power dissipation, traffic assignment
- Lagrange multipliers and Karush Kuhn Tucker conditions; duality
- Wardrop equilibrium: user optimum
- User optimum vs social optimum: price of anarchy!
  Braess paradox
- Toll design