Decentralized control of complex systems: an overview

Madhu N. Belur
(Course taught with Debraj Chakraborty)
Control and Computing group
Department of Electrical Engineering
Indian Institute of Technology Bombay
www.ee.iitb.ac.in/∼belur (for this talk's pdf-file)

VESIT, 10th July, 2014

Outline

- What, why: decentralized control
- Graph theory
- Markov chains, row-stochastic matrices
- M matrices
- Structural controllability
- Matroids
- Multi-agent systems
- Consensus, formation-control

What

- Large systems are increasingly complex
- Interconnection of many simpler subsystems
- Individual subsystems have local controllers
- Each subsystem
  - has (local) actuator inputs
  - has (local) sensor measurements
  - interacts with a 'few' other subsystems (neighbours)
  - allows local controllers
- Two options
  - Central controller accesses all measurements and actuates all actuator inputs
    (Delhi decides Matunga garbage disposal truck route)
  - Local controller (Muncipality decides on local issues)
- Centralized controller: could be more optimal
- Decentralized controller: more effective

Centralized vs decentralized controller

- Complexity of large systems
- Centralized controller requires too much communication
- Communication requirement: bandwidth, delays, reliability
- Decentralized control:
  Settle for (in general) sub-optimal performance, but far more reliable.
- Reliable because: less need to communicate, controllers spread out across large system
- Aim for robustness to absence/presence of interactions (changing graphs)

- Each subsystem's controller can access only local system's variables
- Structure very important, in addition to system parameters
- Computation of controller easier. More 'scalable'
- For example: microgrid: 'smart-grids'

Microgrid/smartgrid

- For an island, often no (large) grid reaches out to give power
- Wind energy, tidal power, solar power, battery and diesel generator connected together to form a 'micro-grid'
- Each DC source has a Voltage-Source-Inverter
- Resulting AC sources are connected to form a grid
- AC frequency (initially) set to 50Hz for each source
- But two AC sources having frequency 50.001 and 50.000000001 will eventually go out of phase.
- Frequency control: possible by 'centralized controller' which measures phase angle of all sources
- These days: time-stamped data from GPS: for synchronization

Decentralized frequency control

Alternatively:

- Each source ‘regulates’ its frequency decentrally so that all sources converge to the same value
- Decentralized frequency control
- Possible? Can a local controller ensure global convergence of frequency?
- Large grids, generators rarely go unstable

Roof top PV panels

With rise in renewable interest:

- Many countries encourage individual houses to ‘tap’ solar power and ‘feed up’ into the grid
- Each house can also ‘generate’ power now
- Frequency control:

Roof top PV panels

With rise in renewable interest:

- Many countries encourage individual houses to 'tap' solar power and 'feed up' into the grid
- Each house can also 'generate' power now
- Frequency control: pricing issues
- Concern about synchronized ON/OFF switching due to pricing policies
- Frequency control law set by Government (Germany)

Droop law: microgrid

- Consider island with a microgrid comprised of several AC sources
- Need to settle to a common frequency value:

Droop law: microgrid

- Consider island with a microgrid comprised of several AC sources
- Need to settle to a common frequency value: quickly
- Most common law: frequency decreases slightly when power drawn exceeds rated power
- Mimic generators (large inertia)
- Linear law: frequency 'droop' proportional to increase in power drawn
- Droop law: results in stability for small droop shown first by Chandorkar, Divan and Adapa (1991):

Droop law: microgrid

- Consider island with a microgrid comprised of several AC sources
- Need to settle to a common frequency value: quickly
- Most common law: frequency decreases slightly when power drawn exceeds rated power
- Mimic generators (large inertia)
- Linear law: frequency 'droop' proportional to increase in power drawn
- Droop law: results in stability for small droop shown first by Chandorkar, Divan and Adapa (1991):
  for ring and radial mesh

Droop law: microgrid

- Consider island with a microgrid comprised of several AC sources
- Need to settle to a common frequency value: quickly
- Most common law: frequency decreases slightly when power drawn exceeds rated power
- Mimic generators (large inertia)
- Linear law: frequency 'droop' proportional to increase in power drawn
- Droop law: results in stability for small droop shown first by Chandorkar, Divan and Adapa (1991):
  for ring and radial mesh
- In 2011: for arbitrary graphs: Iyer, Belur and Chandorkar
- Laplacian matrix plays a role here too!

Where else?

Laplacian matrix: central to most multi-agent systems' studies.

Laplacian Matrix

- Consider an undirected unweighted graph G with vertices v_1, v_2, … v_n, and edges E.
- Define $D \in \mathbb{R}^{n\times n}$ a diagonal 'degree' matrix: d_{ii} is the degree of v_i.
- $A \in \mathbb{R}^{n\times n}$ is the adjacency matrix: $a_{ij} = 1$ if v_i and v_j are adjacent (there is an edge between them).
- Laplacian matrix $L := D - A$.
- $L \geqslant 0$ (non-negative definite matrix)
  $x^T L x \geqslant 0$ for all vectors $x \in \mathbb{R}^n$
- $L \geqslant 0$ (non-negative definite matrix)
  $x^T L x \geqslant 0$ for all vectors $x \in \mathbb{R}^n$

Laplacian of a graph

- # eigenvalues at the origin: # connected components of the graph
- For connected graph, second smallest eigenvalue: algebraic connectivity ν of the graph
- Graph connected $\Leftrightarrow \nu > 0$
- Corresponding eigenvector: zero/nonzero structure very informative:
  S. Pati and R.B. Bapat
- Adding edges can only increase ν: Fiedler, 1973
  Hence 'connectivity'
- Other notions of graph connectivity: computationally hard NP-hard, etc.

Laplacian of a graph

- Connects to Markov chains (stochastic matrices)
- Graph theory, combinatorial optimization
- Matrix theory (M matrices, Hicks, positive matrices, Metzler,
- Perron Frobenius theorem
- Flocking matrix, Diffusion problems of discrete time
- Note: $L := D - A$. But flocking matrix: $D^{-1}A$: a stochastic matrix (Vicsek model)
- Many diffusion problems in discrete time: stochastic matrices
- In multi-agent systems: information 'diffuses' (through neighbours)

Multi-agent systems

- Network of identical agents
- Can communicate with immediate neighbours
- May or may not have a 'leader'
- Leader (if any): some global communication, rest $\equiv$ 'followers'
- Multi-agents need to collectively achieve a task: surveillance.
- Computing also done distributedly: graph connectivity calculation local to get an estimate of algebraic connectivity (global parameter).
- Stability, consensus, formation control
- In spite of changing graphs (as long as some connectivity remains across time-spans.)

A. Jadbabaie, Olfati-Saber, Ali Saberi

Markov chains

- Non-negative square matrices
- Each row adds up to one: row-stochastic matrices
- $A \in \mathbb{R}^{n \times n}$: 'probabilities transition matrix'
- $p(k)$: (column) vector of probabilities (of being in state i) at time k, then
$$\boldsymbol{p^T(k+1) := p^T(k)A}$$
is the vector of probabilities at the next instant $k+1$.
- For A, entries in each row add up to one $\equiv [1\ 1\ \cdots\ 1]^T$ is an eigenvector corresponding to eigenvalue 1.

‘Irreducible’ Markov chains

- Matrix A gives an irreducible Markov chain if there is some nonzero probability of reaching each state in finite time starting in any state at $k = 0$
- The corresponding directed graph is strongly connected
- Perron Frobenius (PF) Theorem gives useful implications for this case
- Perron Frobenius Theorem: applicable for non-negative matrices
- Non-negative matrix: each entry of the matrix is non-negative
- Non-negative matrix $\neq$ Non-negative definite (semidefinite)

Perron Frobenius Theorem

For a matrix $A \in \mathbb{R}^{n \times n}$, with all entries non-negative ($A \geqslant 0$)

- there exists an eigenvalue $\lambda_P > 0$ (unless $A = 0$) with eigenvector v_P having non-negative components ($v_P \geqslant 0$).

Further, assume $A > 0$.

- Then all components in that eigenvector are positive.
- Every other eigenvector has ‘mixed signs’
- ‘Farthest’ eigenvalue (spectral radius) $= \lambda_P$
- Algebraic multiplicity of λ_P is one.
- Take any $v \in \mathbb{R}^n$ with $v \geqslant 0$ (and nonzero).
  Then $A^k v$ ‘turns towards’ v_P as $k \to \infty$.

Google Page rank Algorithm (after Larry Page):
very quick for LARGE sparse matrices
Ranks web-pages based on a very quick sparse matrix calculation

Perron Frobenius theorem: non-negative

$A > 0$: not sparse

$A \geqslant 0$: can be sparse

Close link between directed graph G_A properties and matrix properties of A

Irreducible : strongly connected

Consider A as a (row) stochastic matrix

- Primitive (A^k is positive for some finite k)
- Ergodic ($\lim_{k\to\infty} A^k$ is rank one)

Not primitive: some loops (repeated visits with probability one)

Non-negative matrices: closely linked with M matrix

M matrices

- Square matrices such that off-diagonal entries are nonpositive and
  diagonal entries are 'quite' positive
- A with nonpositive off-diagonal entries is called an M matrix if $A = \rho I - B$
  for a non-negative matrix B and $\rho >$ spectral radius of B.
- A with nonpositive off-diagonal entries is an M matrix if and only if A^{-1} exists and is $\geqslant 0$.
- If A is an M matrix, then $-A$ is Hurwitz.

A is an M matrix

- Interpret $\frac{\mathrm{d}}{\mathrm{d}t}x = -Ax$ as many stable subsystems (diagonal terms),
  with destabilizing 'neighbour' subsystems
  This means diagonal term in $-A$ is negative, and off-diagonal is positive
- Suppose: local subsystem 'very' stable compared to destabilizing neighbour interactions
- Competitive systems
- Stability inspite of competing/destabilizing neighbours
- Siljak: Robust interconnection stability

Markov chains

- Weighted graphs
- Mixing time of Markov chains: rapidly mixing
- Directed graphs, strongly connected graphs.
- Bipartite graphs: matching theory.

Distance matrix

Work by Bapat and Pati (more for Resistance distance matrix)
Conditionally positive definite matrices
Consider an undirected graph with vertices $v_1, \ldots, v_n$.

- Distance matrix D: $d(i, j)$ is minimum number of edges to use to reach j from i
- Consider for just connected graphs (all entries finite)
- Symmetric matrix (for undirected graphs)
- All diagonal elements zero
- Triangular inequality
- OK to use the word distance

Resistance distance matrix: electrical analogy

Resistance distance matrix

Conditionally positive definite matrices
Consider an undirected graph with vertices $v_1, \ldots, v_n$.

- Resistance distance matrix R: $r(i, j)$ is effective resistance between nodes i and j
- Consider again for just connected graphs (all entries finite)
- Symmetric matrix (for undirected graphs)
- All diagonal elements zero
- Triangular inequality: Bapat
- Again OK to use the word distance

Resistance distance matrix: closely linked to Laplacian matrix
Boyd and Arpita Ghosh: use R to identify 'best' edge to add (for algebraic connectivity improvement)

Conditionally positive definite: symmetric matrix, at most one negative eigenvalue, rest all positive.
(positive/non-negative relaxations vary)
L: Laplacian

- e^{-Lt}: row-stochastic
  (for symmetric L and A)
- Connected graph, e^{-Lt} is ergodic

Much work on directed graphs (and corresponding unsymmetric A and L matrices)
Doubly Stochastic matrices:

Non-negative matrices in which both rows and columns add up to one. A convex set: on boundaries: permutation matrices
Good link between convexity properties of set of doubly stochastic matrices
and Hamiltonian cycle problem (NP-hard):

Vivek Borkar, Vladimir Ejov, Jerzy A. Filar and Giang T. Nguyen (Approximation algorithms: polynomial time)

Rapidly mixing Markov Chain

- Second farthest eigenvalue λ_2 (say) of irreducible, aperiodic Markov chain
- Relates to the time-constant as arbitrary initial conditions reach steady-state (stationary) distribution
- Analogous to second smallest eigenvalue ν of Laplacian (for connected graph): inverse of time-constant of network
- Better connected $\equiv$ Faster mixing
  $\equiv$ smaller $|\lambda_2|$ (Markov chain) $\equiv$ higher ν (Laplacian).
- Depending on the context, both are 'spectral gap': ν and $1 - |\lambda_2|$.
  (For stochastic matrix, $\lambda_P = 1$ and for Laplacian matrix, $\lambda_n = 0$.)

Boyd, Diaconis, Xiao: fastest mixing Markov chain: convexify the problem

Matroid techniques

Murota and van der Woude: structural controllability

- Matroid: generalizes key properties from linear algebra (vectors, independence) and graphs (trees, absence of loops)
- M (a set of various subsets of a set E) is a matroid if the objects in M satisfy following properties:
  1. Empty set is in M
  2. If T_1 is in M, then every subset of T_1 is also in M
  3. If T_1 and T_2 are in M and $|T_1| < |T_2|$, then there is an element t of T_2 such that $T_1 \cup t$ also is in M.
- Think of a graph with edge set E and as elements of M, think of all sets of edges that do not contain a loop. (Elements of M are Trees/forests)
- Elements of M are called 'independent sets'.

Structural controllability

- Siljak, van-der-Woude, Murota
- Primarily state-space
- Bipartite graph: matching technique
- Rachel Kalaimani, Belur and Sivaramakrishnan: pole placement

Structural controllability

Key contributors:

Dragoslav D. Šiljak
M. Ikeda, M.E. Sezer
A.S. Morse

More recently,

A. Jadbabaie, Olfati-Saber, Ali Saberi
J. Alexander Fax and Richard M. Murray
Vicsek model (flocking matrices)

More topics

Primitive matrices, disagreement dynamics,
generic: almost any: thin set

Generic statements

Take randomly entries from $\mathbb{R}$ to build a matrix $A \in \mathbb{R}^{n\times n}$
Generically, A is nonsingular
Generically, matrices (A, B) is a controllable pair

For example:
One cart: with two inverted pendulums
controlled by one force: controllable for different lengths
Unlikely, that lengths are exactly equal

Other works

Event triggered computing (in multi-agent systems)

Mesbahi and Zelazo: edge agreement: process noise within each agent, and measurement noise at each edge: performance limitations due to graph constraints

'Manufacturing consent': Vivek Borkar, EE, IIT Bombay

(Noam Chomsky and Hermann have a similarly titled book)

Laplacian: $-L$ is a (singular) M matrix

e^{-Lt} is a flocking matrix (for any time t)

Laplacian: $-L$ is a (singular) M matrix

e^{-Lt} is a flocking matrix (for any time t)
Discretize a continuous LTI system
(Continuous system $\rightarrow$ discrete time system)

Laplacian: $-L$ is a (singular) M matrix

e^{-Lt} is a flocking matrix (for any time t)
Discretize a continuous LTI system
(Continuous system $\rightarrow$ discrete time system)
$\lim_{t\rightarrow\infty} e^{-Lt}$: rank one matrix $\Leftrightarrow$ (undirected) graph is connected

Laplacian: $-L$ is a (singular) M matrix

e^{-Lt} is a flocking matrix (for any time t)
Discretize a continuous LTI system
(Continuous system $\rightarrow$ discrete time system)
$\lim_{t\rightarrow\infty} e^{-Lt}$: rank one matrix $\Leftrightarrow$ (undirected) graph is connected

Ergodic
Siljak: Form I and Form II

Conclusion

- Decentralized control: frequency control
- Active power for frequency,
- Reactive power for Voltage magnitude control
- Graph theory, matroids
- Markov chains, row-stochastic matrices
- M matrices
- Structural controllability: generic controllability
- Consensus, formation-control of multi-agent systems

Conclusion

- Decentralized control: frequency control
- Active power for frequency,
- Reactive power for Voltage magnitude control
- Graph theory, matroids
- Markov chains, row-stochastic matrices
- M matrices
- Structural controllability: generic controllability
- Consensus, formation-control of multi-agent systems

Thank you