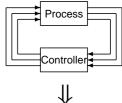
- Introduction to distributed control
- Dual decomposition
- Distributed MPC

 x_1

C

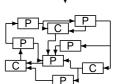
Gradient methods for large-scale systems



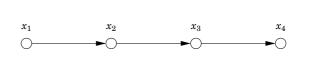


We need methodology for

- Decentralized specifications
- Decentralized design
- Verification of global behavior



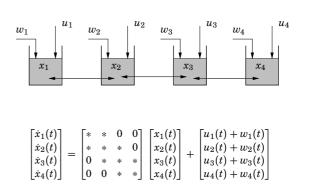
Example 2: A supply chain for fresh products



Fresh products degrade with time:

	$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} =$	* 0 0 0	* 0	0 *	0 0	$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} +$	$\begin{bmatrix} -u_1(t) + w_1(t) \\ u_1(t) - u_2(t) \\ u_2(t) - u_3(t) \\ u_3(t) + w_4(t) \end{bmatrix}$
--	--	------------------	--------	--------	--------	--	---

Example 4: Irrigation Channels



Lecture 4

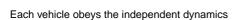
- Introduction to distributed control
- Dual decomposition
- Distributed MPC
- Gradient methods for large-scale systems

 x_2 x_3

 x_4

-0

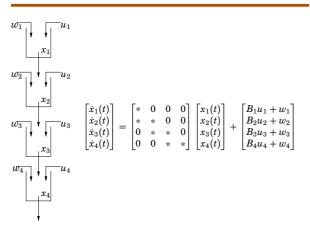
Example 1: A vehicle formation



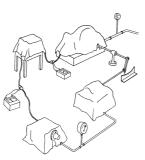
$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} =$	$\begin{bmatrix} * \\ 0 \\ 0 \\ 0 \end{bmatrix}$	0 * 0 0	0 0 * 0	0 0 0 *	$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} +$	$\begin{bmatrix} B_{1}u_{1}(t) + w_{1}(t) \\ B_{2}u_{2}(t) + w_{2}(t) \\ B_{3}u_{3}(t) + w_{3}(t) \\ B_{4}u_{4}(t) + w_{4}(t) \end{bmatrix}$
$\lfloor \dot{x}_4(t) \rfloor$	[0	0	0	*	$\lfloor x_4(t) \rfloor$	$\left\lfloor B_4 u_4(t) + w_4(t) \right\rfloor$

The objective is to make $\mathbf{E}|Cx_{i+1} - Cx_i|^2$ small for $i = 1, \dots, 4$.

Example 3: River Water Flow



Can Systems be Certified Distributively?



Componentwise performance verification without global model?

50 year old idea: Dual decomposition

$\min_{z} [V_1(z_1, z_2) + V_2(z_2) + V_3(z_3, z_2)]$

 $= \max_{p_1} \min_{z_2, v_1} \left| V_1(z_1, v_1) + V_2(z_2) + V_3(z_3, v_3) + p_1(z_2 - v_1) + p_3(z_2 - v_3) \right|$

The optimum is a Nash equilibrium of the following game:

The three computers try to minimize their respective costs Computer 1: $\min_{z_1,v_1} [V_1(z_1,v_1) - p_1v_1]$ Computer 2: $\min_{z_2} [V_2(z_2) + (p_1 + p_3)z_2]$

Computer 3: $\min_{z_3,v_3} [V_3(z_3,v_3) - p_3v_3]$

while the "market makers" try to maximize their payoffs

Between computer 1 and 2: $\max_{p_1} [p_1(z_2 - v_1)]$ Between computer 2 and 3: $\max_{p_3} [p_3(z_2 - v_3)]$

Global stability of saddle algorithm

$$\min_{Rx=0} V(x) = \max_{p} \min_{x} [V(x) + p^{T}Rx]$$

$$\begin{cases} \dot{x} = -G \left[(\partial V / \partial x)^{T} + R^{T}p \right] \\ \dot{p} = HRx \end{cases} \qquad G, H > 0 \text{ adjustment rates}$$

$$\begin{bmatrix} \ddot{x} \\ \ddot{p} \end{bmatrix} = \begin{bmatrix} -G (\partial^{2}V / \partial x^{2}) & -GR^{T} \\ HR & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \ddot{p} \end{bmatrix}$$

$$\mathbf{V} = |\dot{x}|_{G^{-1}}^{2} + |\dot{p}|_{H^{-1}}^{2}$$

$$\frac{d}{d} \mathbf{V} = \dot{x}^{T} G^{-1} \ddot{x} + \dot{p}^{T} H^{-1} \ddot{p}$$

$$\begin{split} \overline{dt} \mathbf{V} &= \dot{x}^* \, G^{-i} \ddot{x} + \dot{p}^T H^{-1} \ddot{p} \\ &= -\dot{x}^T \left[(\partial^2 V / \partial x^2) \dot{x} + R^T \dot{p} \right] + \dot{p}^T (R \dot{x}) \\ &= -\dot{x}^T (\partial^2 V / \partial x^2) \dot{x} \leq 0 \end{split}$$

Important Aspects of Dual Decomposition

- Very weak assumptions on graph
- No need for central coordination
- Natural learning procedure is globally convergent
- Unique Nash equilibrium corresponds to global optimum

Conclusion: Ideal for control synthesis by prescriptive games

Case study: A water supply network in Paris

[Carpentier and Cohen, Automatica 1993]

- Network serving about 1 million inhabitants
- 20 main water reservoirs
- 30 pumping stations
- 13 peripheral subnetworks

Challenges for control

- Minimize cost for pumping
- Bounds on reservoirs
- Bounds and delays in pumping power
- Prediction of consumption

Optimal control using dual decomposition and saddle algorithm Subnetworks separated by two variables: Water flow and price

The saddle point algorithm

Globally convergent if V_i are convex! Lyapunov function: $\sum_i |\dot{z}_i| + \sum_i |\dot{v}_i| + \sum_i |\dot{p}_i|$ [Arrow, Hurwicz, Usawa 1958]

A long history

The saddle algorithm: Arrow, Hurwicz, Usawa 1958

Books on control and coordination in hierarchical systems: Mesarovic, Macko, Takahara 1970 Singh, Titli 1978 Findeisen 1980

Major application to water supply network: Carpentier and Cohen, Automatica 1993

Decentralized Bounds on Suboptimality

Given any $p_1, p_3, \bar{z}_1, \bar{z}_2, \bar{z}_3$, the distributed test

$$\begin{split} &V_1(\bar{z}_1, \bar{z}_2) - p_1 \bar{z}_2 \leq \alpha \min_{z_1, v_1} \left[V_1(z_1, v_1) - p_1 v_1 \right] \\ &V_2(\bar{z}_2) + (p_1 + p_3) \bar{z}_2 \leq \alpha \min_{z_2} \left[V_2(z_2) + (p_1 + p_3) z_2 \right] \\ &V_3(\bar{z}_3, \bar{z}_2) - p_3 \bar{z}_2 \leq \alpha \min_{z_3, v_3} \left[V_3(z_3, v_3) - p_3 v_3 \right] \end{split}$$

implies that the globally optimal cost J^* is bounded as

 $J^* \le V_1(\bar{z}_1, \bar{z}_2) + V_2(\bar{z}_2) + V_3(\bar{z}_3, \bar{z}_2) \le \alpha J^*$

Proof: Add both sides up!

Lecture 4

- Introduction to distributed control
- Dual decomposition
- Distributed MPC
- Gradient methods for large-scale systems

A control problem with graph structure

$$\begin{bmatrix} x_1(\tau+1)\\ x_2(\tau+1)\\ \vdots\\ x_J(\tau+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & 0\\ A_{21} & \ddots & \ddots\\ & \ddots & \ddots\\ 0 & A_{J(J-1)} & A_{JJ} \end{bmatrix} \begin{bmatrix} x_1(\tau)\\ x_2(\tau)\\ \vdots\\ x_J(\tau) \end{bmatrix} + \begin{bmatrix} u_1(\tau)\\ u_2(\tau)\\ \vdots\\ u_J(\tau) \end{bmatrix}$$

Minimize the convex objective $\sum_{t=0}^N \sum_{t=0}^J \ell_i(x_i(\tau), u_i(\tau))$

i=1 $\ell(x(\tau).u(\tau))$

with convex constraints $x_i(\tau) \in X_i$, $u_i(\tau) \in U_i$ and $x(0) = \bar{x}$.

Decomposing the Cost Function

$$\begin{split} & \max_{p} \min_{u,v,x} \sum_{\tau=0}^{N} \sum_{i=1}^{\mathcal{I}} \left[\ell_{i}(x_{i},u_{i}) + p_{i}^{T} \left(v_{i} - \sum_{j \neq i} A_{ij} x_{j} \right) \right] \\ & = \max_{p} \sum_{i} \min_{u_{i},x_{i}} \sum_{\tau=0}^{N} \underbrace{\left[\ell_{i}(x_{i},u_{i}) + p_{i}^{T} v_{i} - x_{i}^{T} \left(\sum_{j \neq i} A_{ji}^{T} p_{j} \right) \right]}_{\ell_{i}^{p}(x_{i},u_{i},v_{i})} \end{split}$$

so, given the sequences $\{p_j(t)\}_{t=0}^N$, agent *i* should minimize

what he expects others to charge him

$$\sum_{\tau=0}^{N} \ell_i(x_i, u_i) + \sum_{\tau=0}^{N} p_i^T v_i - \sum_{\tau=0}^{N} x_i^T \left(\sum_{j \neq i} A_{ji}^T p_j \right)$$

local cost what he is payed by others

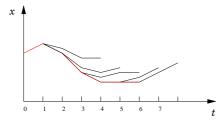
subject to $x_i(t+1) = A_{ii}x_i(t) + v_i(t) + u_i(t)$ and $x_i(0) = \bar{x}_i$.

Idea of Distributed Model Predicitve Control

Replace the original problem by iterative online solutions of the decentralized finite horizon problem

$$\min_{x_i,u_i}\sum_{t=0}^N l_i^p(x_i(t),u_i(t),v_i(t))$$

Two sources of error: Finite horizon and non-optimal prices



Fixed or flexible parameters N_i , K_i , γ_i ?

Fixed parameters

- Simpler implementation
- Gives distributed LTI controllers
- Can be analyzed off-line or on-line

Flexible parameters

- Useful to handle hard state constraints
- Can speed up on-line computations
- Can slow down on-line computations

Decomposing the problem

Minimize
$$\sum_{t=0}^{N} \ell(x(\tau), u(\tau))$$

subject to

$$\begin{bmatrix} x_1(\tau+1)\\ x_2(\tau+1)\\ \vdots\\ x_J(\tau+1) \end{bmatrix} = \begin{bmatrix} A_{11}x_1(\tau)\\ A_{22}x_2(\tau)\\ \vdots\\ A_{JJ}x_J(\tau) \end{bmatrix} + \begin{bmatrix} v_1(\tau)\\ v_2(\tau)\\ \vdots\\ v_J(\tau) \end{bmatrix} + \begin{bmatrix} u_1(\tau)\\ u_2(\tau)\\ \vdots\\ u_J(\tau) \end{bmatrix}$$

where $x(0) = \bar{x}$ and

$$v_i = \sum_{j \neq i} A_{ij} x_j$$

holds for all *i*.

Distributed Optimization Procedure

Local optimizations in each node

$$V_i^{N,p}(\bar{x}_i) = \min_{u_i, x_i} \sum_{\tau=0}^N \ell_i^p(x_i(\tau), u_i(\tau), v_i(\tau))$$

can be coordinated by (local) gradient updates of the prices

$$p_i^{k+1}(\tau) = p_i^k(\tau) + \gamma_i^k \left[v_i^k(\tau) - \sum_{j \neq i} A_{ij} x_j^k(\tau) \right]$$

Future prices included in negotiation for first control input!

Convergence guaranteed under different types of assumptions on the step size sequence γ_i^k .

A Distributed MPC Algorithm

At time t:

- 1. Measure the states $x_i(t)$ locally.
- 2. Use gradient iterations to generate
 - ▶ price prediction sequences $\{p_i(t, \tau)\}_{\tau=0}^N$
 - state prediction sequences $\{x_i(t, \tau)\}$
 - ► input prediction sequences $\{u_i(t, \tau)\}_{\tau=1}^N$

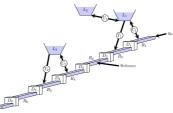
warm-starting from predictions at time t - 1.

3. Apply the inputs $u_i(t) = u_i(t, 0)$.

Important parameters: Prediction horizons N_i , number of gradient iterations K_i and gradient step sizes γ_i .

Hydro Power Valley

Benchmark in EU-project HD-MPC coordinated from Delft





Equipped with 10 turbines $(T_1, T_2, D_1, \dots, D_6, C_{1t}, C_{2t})$ and 2 pumps (C_{1p}, C_{2p}) 3 reservoirs (lakes) 6 dams and reaches

Objectives: Follow power-reference. Avoid flooding.

Modeling:

- 1. Saint Venant PDE (mass and volume balance)
- Spatial discretization (MOL) ⇒ non-linear ODE:s (249 states, 12 inputs, used as simulation model)
- 3. Linearization, discretization (h=30 min) and model reduction ⇒ MPC-model (32 states, 12 inputs)

Difficulties:

- ► Non-linear power-production $p(t) = u(t)^T S_i x(t)$ - Linearize around nominal working point (x_0, u_0) , $\Delta p = u_0^T S \Delta x + x_0^T S^T \Delta u$
- ► Non-linear constraints; $u_{C_{it}}u_{C_{ip}} = 0$, i = 1, 2- We have $u_{C_{it}} \ge 0$ and $u_{C_{ip}} \le 0$, penalize $-u_{C_{it}}u_{C_{ip}}$

Cost function:

$$\sum_{t=0}^{N-1} \left(\frac{1}{2} \begin{bmatrix} \Delta x(t) \\ \Delta u(t) \end{bmatrix}^T H \begin{bmatrix} \Delta x(t) \\ \Delta u(t) \end{bmatrix} + \gamma \left\| P \begin{bmatrix} \Delta x(t) \\ \Delta u(t) \end{bmatrix} - \Delta p_{ref}(t) \right\|_1 \right)$$

with $P = [u_0^T S \ x_0^T S^T]$

Notation

For a distributed accuracy test, let $\bar{V}_i^p(x_i)$ be an upper bound on

$$\min_{u_i,v_i,x_i}\sum_{\tau=0}^{\infty}\ell_i^p(x_i(\tau),u_i(\tau),v_i(\tau))$$

Such an upper bound can for example be computed by minimization over a finite time horizon with a terminal constraint at the origin.

Theorem on accuracy of distributed MPC

Suppose x(t+1) = Ax(t) + Bu(t) for $t \ge 0$ and for some p that

$$\begin{split} & V_i^{N_i(t),p(t,\cdot)}(x_i(t)) + (1-\alpha)\ell_i\big(x_i(t),u_i(t)\big) \\ & \geq \bar{V}_i^{p(t,\cdot)}\big(x_i(t+1)\big) + \ell_i^{p(t,\cdot)}\big(x_i(t),u_i(t),\sum_{j\neq i}A_{ij}x_j(t)\big) \end{split}$$

for all *i* and *t*. Then

$$\alpha \sum_{t=0}^{\infty} \ell(x(t), u(t)) \le V^{\infty}(\bar{x})$$

Notice: Failure of inequality hints on update of N_i or K_i !

Lecture 4

- Introduction to distributed control
- Dual decomposition
- Distributed MPC
- Gradient methods for large-scale systems

Conclusions on Distributed MPC

We have synthesized a game that solves optimal control problems via independent decision-makers in every node, acting in their own interest!

- Optimal strategies independent of global graph structure!
- States are measured only locally
- Linearly complexity (given horizon and iteration scheme)
- Distributed bounds on distance to optimality

Controller Tuning for Large Tri-diagonal Plant

Minimize $V = \mathbf{E} \sum_{i=1}^{n} (|x_i|^2 + |u_i|^2)$

	0.6 0.3	0.1 ·.	·	0	$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ \vdots \\ x_n(t+1) \end{bmatrix} +$	$\begin{bmatrix} u_1(t) + w_1(t) \\ u_2(t) + w_2(t) \\ \vdots \end{bmatrix}$
$\begin{bmatrix} \vdots \\ x_n(t+1) \end{bmatrix}^{-1}$	0	•••	· 0.3	0.1 0.6	$\begin{bmatrix} \vdots \\ x_n(t+1) \end{bmatrix}$	$\begin{bmatrix} \vdots \\ u_n(t) + w_n(t) \end{bmatrix}$

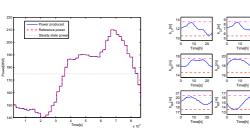
We will optimize a tri-diagonal control structure

$$\bar{L} = \begin{bmatrix} * & * & 0 \\ * & \ddots & \\ & \ddots & * \\ 0 & * & * \end{bmatrix}$$

Figure: Power reference tracking (left) and Dam water levels (right)

Control horizon: N = 10 (5 hours)

Simulation results



Computing the closed loop control performance

We are applying the control law u = -Lx to the system

$$x(t+1) = Ax(t) + Bu(t) + w(t)$$

where w is white noise with variance W. Define

$$J(L) = \mathbf{E} \left(|x|_Q^2 + |u|_R^2 \right)$$

Then the gradient with respect to a particular element L_{ij} is

$$(\nabla_L J)_{ij} = 2RL\mathbf{E} [x_i x_j^T] + 2B^T \mathbf{E} [p_i x_j^T]$$

where p(t) is the stationary solution of the adjoint equation

$$p(t-1) = (A - BL)^T p(t) - (Q + L^T RL)x(t)$$

Gradient iteration for the wind park

A distributed synthesis procedure

- 1. Measure the states $x_i(t)$ for $t = t_k, \ldots, t_k + N$
- 2. Simulate the adjoint equation

$$p_i(t-1) = \sum_{j \in E_i} (A - BL)_{ji}^T p_j(t) - 2(Q_i x_i(t) - \sum_{j \in E_i} L_{ji}^T R_j u_j(t))$$

for $t = t_k, \ldots, t_k + N$ by communicating states between nodes.

3. Calculate the estimates of
$$\mathbf{E} u_i x_j^T$$
 and $\mathbf{E} p_i x_j^T$ by

$$\left(\mathbf{E}\,u_{i}x_{j}^{T}\right)_{\text{est}} = \frac{1}{N+1}\sum_{t=t_{k}}^{t_{k}+N}u_{i}(t)x_{j}(t)^{T} \quad \left(\mathbf{E}\,p_{i}x_{j}^{T}\right)_{\text{est}} = \frac{1}{N+1}\sum_{t=t_{k}}^{t_{k}+N}p_{i}(t)x_{j}(t)^{T}$$

4. For fixed step length $\gamma > 0$, update $L_{ij}^{(k+1)} = L_{ij}^{(k)} + 2\gamma R_i \left(\mathbf{E} u_i x_j^T\right)_{\text{est}} + B_i^T \left(\mathbf{E} p_i x_j^T\right)_{\text{est}}.$ Let $t_{k+1} = t_k + N$ and start over.

Gradient iteration for the wind park

cost	=					cost =					
	14.98	87				10.542	29				
L =						L =					
	0	0	0	0	0	0.0327	0.040	0 0	0	0	
	0	0	0	0	0	-0.0007	0.056	0.0527	0	0	
	0	0	0	0	0	C	-0.006	9 0.0434	0.0315	0	
	0	0	0	0	0	C)	0 -0.0207	0.0131	0.0437	
	0	0	0	0	0	C)	0 C	-0.0033	0.0373	

Gradient iteration for the wind park

cost =					
7.8184					
L =					
0.0310	0.0595	0	0	0	
-0.0168	0.1002	0.1151	0	0	
0	0.0345	0.1357	0.0986	0	
0	0	0.0636	0.0831	0.1351	
0	0	0	0.0102	0.1295	

Gradient iteration for the wind park cost = 7.4004 L = 0.0576 0.0583 0 0 0 1150.12240.138100.03730.1500 0.0115 0 0 0.1153 0 0 0 0.0546 0.1068 0.1566

0

0.0168

0.1594

0

0

Gradient iteration for the wind park

cost =					
7.6192					
L =					
0.0404 -0.0086 0 0 0	0.0685 0.1076 0.0382 0 0	0 0.1193 0.1421 0.0593 0	0 0.1094 0.0991 0.0131	0 0 0.1449 0.1348	

Gradient iteration for the wind park

cost =				
7.2493				
<u>L</u> =				
0.0712	0.0654	0	0	0
0.0061	0.1224	0.1443	0	0
0	0.0341	0.1550	0.1166	0
0	0	0.0773	0.1409	0.1580
0	0	0	0.0418	0.1601

Gradient iteration for the wind park

cost =					
6.9736					
L =					
0.0936	0.1056	0	0	0	
0.0331	0.1775	0.1341	0	0	
0	0.0563	0.1500	0.1215	0	
0	0	0.0700	0.1564	0.1567	
0	0	0	0.0567	0.1646	

cost =					
6.8211					
L =					
0.1390	0.1070	0	0	0	
0.0357	0.1821	0.1549	0	0	
0	0.0668	0.1797	0.1098	0	
0	0	0.0633	0.1685	0.1413	
0	0	0	0.0589	0.1754	

Gradient iteration for the wind park

cost =				
6.7464				
L =				
0.1438	0.1208	0	0	0
0.0470	0.2031	0.1632	0	0
0	0.0749	0.1909	0.1046	0
0	0	0.0779	0.1843	0.1388
0	0	0	0.0445	0.1732

Control of a Large Deformable Mirror

Case study of a 1 m diameter deformable mirror, for adaptive optics in large telescopes. Used to correct for aberrations introduced by the atmosphere.

Using finite element method a spatially discretized model.

$$\mathcal{M}\ddot{\boldsymbol{\xi}} + \mathcal{C}\dot{\boldsymbol{\xi}} + \mathcal{K}\boldsymbol{\xi} = \boldsymbol{F}$$

- ▶ 6128 discretization points, each with 6 degrees of freedom.
- 372 force actuators.
- 1136 position sensors.

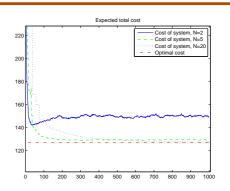
Method data and performance

Method data

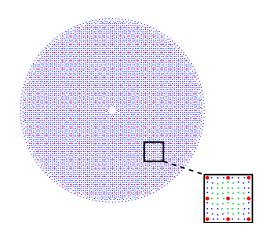
- The sparsity of feedback matrix L is 0.63%.
- Time horizon in gradient computation is 1000 time samples.
- 1000 update iterations are performed.

The computation time for the method becomes 16.6 hours. 70% of this time is spent on calculating matrix inversions in the system simulation.

Performance Versus Number of Gradient Iterations

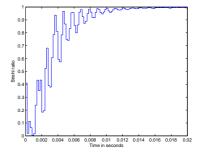


A distributed controller with 100 agents, using only local data. Fewer gradient iterations gives faster convergence, but worse stationary performance.



Control Performance

The controller is used on the mirror when using a simulated atmosphere. Strehl ratio is a common measure in adaptive optics. Defined by $S = e^{-(2\pi\epsilon(t)/\lambda)^2}$ where $\epsilon(t)$ is the RMS error at time *t*.



Lecture Summary

Heavy off-line computations and memory requirements

Dynamic Programming (Explicit MPC)

Extremely fast on-line
 Model Predictive Control

Distributed MPC

Heavy on-line computations

Wide range of applications

Simplifies on-line computations

Reduces communication needs

Gradient methods for large-scale systems

LCCC Focus Periods at Lund University

Invited world-leading researchers from Control, Computer Science, Economics, Communication, Mathematics, ...

In the past:

- Multi-agent coordination and estimation
- Distributed decisions via games and price mechanisms
- Adaptation and learning in autonomous systems
- Distributed model predictive control and supply chains

Upcoming:

- System design meets equation-based languages (September 19-21, 2012)
- Information and control in networks (October 2012)
- Formal verification of embedded control systems (April-13)

See www.lccc.lth.se and announcements.