Distributed Robustness Analysis

Anders Hansson

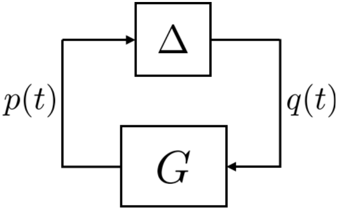
Division of Automatic Control
Linköping University

June 16, 2017

Outline

Robustness Analysis

Consider the following uncertain system,



$$p = Gq, \quad q = \Delta(p), \tag{1}$$

where $G \in \mathcal{RH}_\infty^{p\times m}$ is a transfer function matrix, and $\Delta : \mathcal{L}_2^p \to \mathcal{L}_2^m$ is a bounded and causal operator.

The uncertain system in (1) is said to be robustly stable if the interconnection between G and Δ remains stable for all Δ in some class.

Integral Quadratic Constraints

Let $\Delta : \mathcal{L}_2^p \to \mathcal{L}_2^m$ be a bounded and causal operator. This operator is said to satisfy the IQC defined by Π, i.e., $\Delta \in \mathrm{IQC}(\Pi)$, if

$$\int_0^\infty \begin{bmatrix} v \\ \Delta(v) \end{bmatrix}^T \Pi \begin{bmatrix} v \\ \Delta(v) \end{bmatrix} dt \geq 0, \quad \forall v \in \mathcal{L}_2^p \ , \qquad (2)$$

where Π is a bounded and self-adjoint operator. Assuming that Π is linear time-invariant and has a transfer function matrix representation, the IQC in (2) can be written in the frequency domain as

$$\int_{-\infty}^\infty \begin{bmatrix} \widehat{v}(j\omega) \\ \widehat{\Delta(v)}(j\omega) \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} \widehat{v}(j\omega) \\ \widehat{\Delta(v)}(j\omega) \end{bmatrix} d\omega \geq 0, \qquad (3)$$

where $\hat{v}$ and $\widehat{\Delta(v)}$ are the Fourier transforms of the signals

Stability Theorem

Theorem (IQC analysis)

The uncertain system in (1) *is robustly stable, if*

1. *for all* $\tau \in [0,1]$ *the interconnection described in* (1), *with* $\tau\Delta$, *is well-posed;*
2. *for all* $\tau \in [0,1]$, $\tau\Delta \in \mathrm{IQC}(\Pi)$*;*
3. *there exists* $\epsilon > 0$ *such that*

$$\begin{bmatrix} G(j\omega) \\ I \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} G(j\omega) \\ I \end{bmatrix} \preceq -\epsilon I, \quad \forall \omega \in [0, \infty]. \tag{4}$$

Proof.

See Megretski and Rantzer, 1997. □

Example

If Δ is a linear operator, i.e. $q = \Delta p$, where $\Delta = \delta I$, $\delta \in [-1, 1]$, then

$$\Pi(j\omega) = \begin{bmatrix} X(j\omega) & Y(j\omega) \\ Y(j\omega)^* & -X(j\omega) \end{bmatrix}$$

where $X(j\omega) = X(j\omega)^* \succeq 0$ and $Y(j\omega) = -Y(j\omega)^*$.

Typically Π is parameterized with basis functions.

Collection of Uncertain Systems

Consider a collection of uncertain systems:

$$\begin{aligned} p^i &= G^i_{pq} q^i + G^i_{pw} w^i \\ z^i &= G^i_{zq} q^i + G^i_{zw} w^i \\ q^i &= \Delta^i(p^i), \end{aligned} \tag{5}$$

and let $p = (p^1, \ldots, p^N)$, $q = (q^1, \ldots, q^N)$, $w = (w^1, \ldots, w^N)$ and $z = (z^1, \ldots, z^N)$.

Interconnection of Uncertain Systems

$$\underbrace{\begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^N \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \cdots & \Gamma_{1N} \\ \Gamma_{21} & \Gamma_{22} & \cdots & \Gamma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{N1} & \Gamma_{N2} & \cdots & \Gamma_{NN} \end{bmatrix}}_{\Gamma} \underbrace{\begin{bmatrix} z^1 \\ z^2 \\ \vdots \\ z^N \end{bmatrix}}_{z} \tag{6}$$

Each of the blocks Γ_{ij} are 0-1 matrices.

Interconnected uncertain system:

$$\begin{aligned} p &= G_{pq} q + G_{pw} w \\ z &= G_{zq} q + G_{zw} w \\ q &= \Delta(p) \\ w &= \Gamma z, \end{aligned} \tag{7}$$

where $G_{\star\bullet} = \mathrm{diag}(G^1_{\star\bullet}, \ldots, G^N_{\star\bullet})$ and $\Delta = \mathrm{diag}(\Delta^1, \ldots, \Delta^N)$.

Lumped Formulation

Eliminate w:

$$p = \bar{G}q, \quad q = \Delta(p), \tag{8}$$

where $\bar{G} = G_{pq} + G_{pw}(I - \Gamma G_{zw})^{-1}\Gamma G_{zq}$.

The interconnected uncertain system is robustly stable if there exists a matrix $\bar{\Pi}$ such that

$$\begin{bmatrix} \bar{G}(j\omega) \\ I \end{bmatrix}^* \bar{\Pi}(j\omega) \begin{bmatrix} \bar{G}(j\omega) \\ I \end{bmatrix} \preceq -\epsilon I, \quad \forall \omega \in [0, \infty], \tag{9}$$

for some $\epsilon > 0$. LMI is *dense*.

Sparse Formulation

Theorem

Let $\Delta \in \mathrm{IQC}(\bar{\Pi})$. If there exist $\bar{\Pi}$ and $X = xI \succ 0$ such that

$$\begin{bmatrix} G_{pq} & G_{pw} \\ G_{zq} & G_{zw} \\ I & 0 \\ 0 & I \end{bmatrix}^* \begin{bmatrix} \bar{\Pi}_{11} & 0 & \bar{\Pi}_{12} & 0 \\ 0 & -\Gamma^T X \Gamma & 0 & \Gamma^T X \\ \bar{\Pi}_{21} & 0 & \bar{\Pi}_{22} & 0 \\ 0 & X\Gamma & 0 & -X \end{bmatrix} \begin{bmatrix} G_{pq} & G_{pw} \\ G_{zq} & G_{zw} \\ I & 0 \\ 0 & I \end{bmatrix} \preceq -\epsilon I, \quad (10)$$

for $\epsilon > 0$ and for all $\omega \in [0, \infty]$, then the interconnected uncertain system in (7) is robustly stable.

Sparsity in SDPs

General SDP (new definition of x):

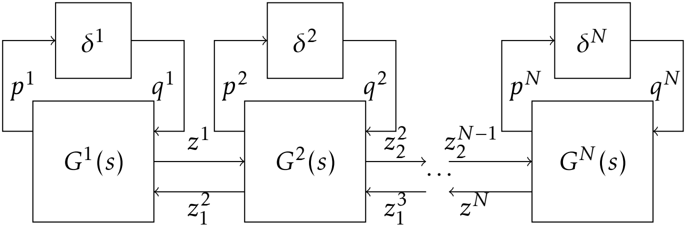
$$\begin{aligned} \underset{S,x}{\text{minimize}} \quad & c^T x && \text{(11a)} \\ \text{subject to} \quad & F^0 + \sum_{i=1}^{m} x_i F^i + S = 0, \quad S \succeq 0. && \text{(11b)} \end{aligned}$$

with $S \in \mathbf{S}^n$, $x \in \mathbb{R}^m$, $c \in \mathbb{R}^m$ and $F^i \in \mathbf{S}^n$ for $i = 0, \ldots, m$.

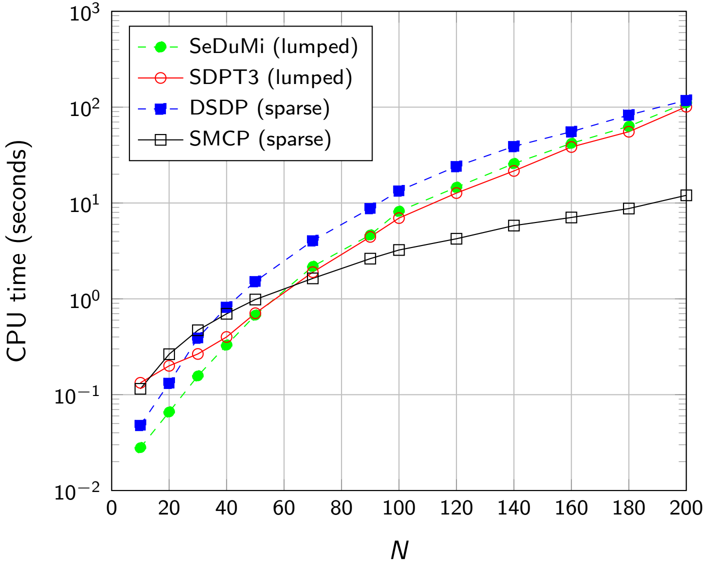
Slack variable S inherits sparsity pattern from problem data.

Solvers like DSDP (Benson and Ye, 2005) and and SMCP (Andersen, Dahl and Vandenberghe, 2010) make use of this structure.

Chain of Uncertain Systems



Average CPU Time



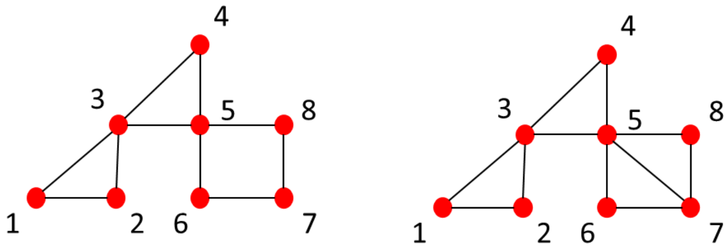
Sparsity Graph

A *sparsity pattern* is a set $E \subseteq \{\{i, j\} \mid i, j \in \{1, 2, \dots, n\}\}$.

A matrix $A \in \mathbf{S}^n$ is said to have a sparsity pattern E if $A_{i,j} = A_{j,i} = 0$, whenever $i \neq j$ and $\{i, j\} \notin E$, or equivalently $A \in \mathbf{S}^n_E$.

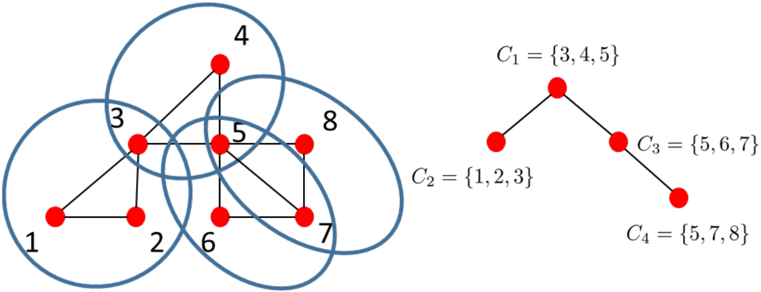
The graph $G = (V, E)$ with $V = \{1, 2, \dots, n\}$ is called the *sparsity graph* associated with the sparsity pattern.

Chordal Graphs and Sparsity Patterns



$$A = \begin{bmatrix} x & x & x & 0 & 0 & 0 & 0 & 0 \\ x & x & x & 0 & 0 & 0 & 0 & 0 \\ x & x & x & x & x & 0 & 0 & 0 \\ 0 & 0 & x & x & x & 0 & 0 & 0 \\ 0 & 0 & x & x & x & x & * & x \\ 0 & 0 & 0 & 0 & x & x & x & 0 \\ 0 & 0 & 0 & 0 & * & x & x & x \\ 0 & 0 & 0 & 0 & x & 0 & x & x \end{bmatrix}$$

Cliques and Clique Trees



A *maximal clique* C_i is a maximal subset of V such that its induced subgraph is complete.

A tree of maximal cliques for which $C_i \cap C_j$ for $i \neq j$ is contained in all the cliques on the path connecting C_i and C_j is said to have the *clique intersection property*. (Always exists.)

Sparse Cholesky Factorization

A sparsity pattern E is chordal if and only if any positive definite matrix $A \in \mathbf{S}_E^n$ has a Cholesky factorization $PAP^T = LDL^T$ with

$$P^T(L + L^T)P \in S_E^n$$

for some permutation matrix P, which is related to the clique intersection property.

After permutation sparse postive definite matrices with chordal sparsity pattern have sparse Cholesky factorizations with no fill-in.

Test for Positive Semidefiniteness (Grone et al., 1984)

A partially specified matrix $A \in \mathbf{S}^n$ can be completed to a positive semidefinite matrix if and only if

$$A_{C_i} \succeq 0$$

where C_i are the maximal cliques of the graph for the specified entries. (A_{C_i} denotes the sub-matrices obtained by picking out the columns and rows indexed by C_i)

Example:

$$\begin{bmatrix} 1 & 1/2 & ? \\ 1/2 & 1 & 1/3 \\ ? & 1/3 & 1 \end{bmatrix} \succeq 0 \Leftrightarrow \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \succeq 0 \ \& \ \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1 \end{bmatrix} \succeq 0$$

Dual SDP

Primal problem again:

$$\underset{S,x}{\text{minimize}} \quad c^T x \tag{12a}$$

$$\text{subject to} \quad F^0 + \sum_{i=1}^{m} x_i F^i + S = 0, \quad S \succeq 0. \tag{12b}$$

with chordal S with cliques C_j, $j = 1, \ldots, p$.

Dual SDP:

$$\underset{Z}{\text{minimize}} \ \operatorname{tr} ZF^0 \tag{13a}$$

$$\text{subject to} \ \operatorname{tr} ZF^i = c_i, \ \ i = 1, \ldots, m \tag{13b}$$

$$Z \succeq 0 \tag{13c}$$

Domain-Space Decomposition (Fukuda et al., 2000)

Write $F^i = \sum_{j=1}^{p} E_j F_j^i E_j^T$ with E_j containing columns of identity matrix indexed by clique C_j. (Not unique)

Since $\operatorname{tr} ZF^i = \sum_{j=1}^{p} \operatorname{tr} E_j^T Z E_j F_j^i$, equivalent dual problem is:

$$\underset{Z}{\text{minimize}} \ \sum_{j=1}^{p} \operatorname{tr} Z_{C_j} F_j^0 \tag{14a}$$

$$\text{subject to} \ \sum_{j=1}^{p} \operatorname{tr} Z_{C_j} F_j^i = c_i, \ \ i = 1, \ldots, m \tag{14b}$$

$$Z_{C_j} \succeq 0 \ \ i = 1, \ldots, p \tag{14c}$$

Consensus Constraints

Equivalantly in decoupeled form:

$$\underset{Z}{\text{minimize}} \sum_{j=1}^{p} \operatorname{tr} Z_j F_j^0 \tag{15a}$$

$$\text{subject to} \sum_{j=1}^{p} \operatorname{tr} Z_j F_j^i = c_i, \ \ i = 1, \ldots, m \tag{15b}$$

$$Z_j \succeq 0, \ \ j = 1, \ldots, p \tag{15c}$$

$$E_{i,j}^T \left(E_i Z_i E_i^T - E_j Z_j E_j^T \right) E_{i,j} = 0, \tag{15d}$$

$\forall\ i, j$, where i are children of j in a clique tree with the clique intersection property, and where j are all non-leaf nodes of the tree. $E_{i,j}$ contains the columns of the identity matrix indexed by $C_i \cap C_j$.

Range-Space Decomposition (Fukuda et al., 2000)

The dual of the previous problem is

$$\underset{x,U}{\text{minimize}}\ c^T x \tag{16a}$$

$$\text{subject to } F_j^0 + \sum_{i=1}^{m} x_i F_j^i + G_j(U) \succeq 0 \ j = 1, \ldots, p \tag{16b}$$

where $x \in \mathbb{R}^m$, with

$$G_j(U) = E_k^T E_{k,j} U_{k,j} E_{k,j}^T E_k - \sum_{i \in \mathrm{ch}(j)} E_j^T E_{i,j} U_{i,j} E_{i,j}^T E_j$$

where $U_{i,j} \in \mathbf{S}^{|C_i \cap C_j|}$, and where k is the parent of j in the clique tree. (For the root and for the leafs some of the terms are not there)

Often the above LMIs are loosely coupled, i.e. many F_j^i are zero.

Example

Find $x = (x_1, \ldots, x_4)$ such that

$$\begin{bmatrix} x_1 & x_2 & 0 \\ x_2 & x_1 & x_3 \\ 0 & x_3 & x_4 \end{bmatrix} \succeq 0$$

is equivalent to find (x, u) such that

$$\begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 + u \end{bmatrix} \succeq 0 \ \& \ \begin{bmatrix} -u & x_3 \\ x_3 & x_4 \end{bmatrix} \succeq 0$$

Can derive distributed Alternating Linearization Method (ALM) on equivalent formulation of feasibility problem. (Details skipped)

Scale-Free Network

- Interconection of 500 subsystems over randomly generated scale-free network, in this case a tree.
  - 478 systems connected to 5 or less other systems
  - 16 systems connected to less than 11 but more than 5 other systems
  - 6 system connected to more than 10 other systems

Scale-Free Network ctd.

- Lumped formulation: LMI of dimension 500 with 500 variables
- Sparse formulation: LMI of dimension 1498 with 1498 variables
- Chordal embedding has 579 cliques with 9894 variables
- Largest LMI has dimension 210 and 170 variables, but 94% of them has dimension 50 or less
- The largest coupling between LMIs involves 92 variables, but 95% of them involve less than 24 variables.
- One of the agents require information from 52 other agents, but 96 % of the agents only require information from at most 10 other agents.

Numerical Results

Solver	Avg. CPU time [sec]
SDPT3 (lumped)	5640
SeDuMi (lumped)	2760
DSDP (sparse)	167
SMCP (sparse)	33
ALM (sparse)	1623

- ALM only prototyped in Matlab
- ALM can use paralell processors
- ALM respect privacy

Domain-Space Decomposition Revisited

Another equivalent formulation of the dual problem is:

$$\underset{Z,Z_j}{\text{minimize}} \sum_{j=1}^{p} \operatorname{tr} Z_j F_j^0 \tag{17a}$$

$$\text{subject to} \sum_{j=1}^{p} \operatorname{tr} Z_j F_j^i = c_i, \ i = 1, \dots, m \tag{17b}$$

$$Z_j \succeq 0, \ j = 1, \dots, p \tag{17c}$$

$$E_j^T Z E_j = Z_j, \ j = 1, \dots, p \tag{17d}$$

where we now have many more variables due to the additional variable Z.

Remember that many F_j^i are zero.

Search Directions for Interior-Point (IP) Methods

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta \bar{z} \\ \Delta \bar{x} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

where H and A are sparse. ($\bar{z}$ vector of all elements of Z and Z_i, $i = 1, 2, \ldots, p$)

Equivalently the optimiality conditions of

$$\underset{\Delta \bar{z}}{\text{minimize}} \; \frac{1}{2} \Delta \bar{z}^T H \Delta \bar{z} - r_1^T \Delta \bar{z} \tag{18a}$$

$$\text{subject to } A \, \Delta \bar{z} = r_2 \tag{18b}$$

After Elimination of the Z_j-variables

$$\underset{\Delta\tilde{z}}{\text{minimize}}\ \frac{1}{2}\Delta\tilde{z}^T\tilde{H}\Delta\tilde{z} - \tilde{r}_1^T\Delta\tilde{z} \tag{19a}$$

$$\text{subject to}\ \tilde{A}\,\Delta\tilde{z} = \tilde{r}_2 \tag{19b}$$

which still has sparse data matrices.

Allmost separable

$$\underset{\Delta z_i}{\text{minimize}}\ \frac{1}{2}\sum_{i=1}^{p}\Delta\tilde{z}_i^T\tilde{H}_i\Delta\tilde{z}_i - \tilde{r}_{1,i}^T\Delta\tilde{z}_i \tag{20a}$$

$$\text{subject to}\ \sum_{j\in\tilde{\mathcal{J}}_i}\tilde{A}_{i,j}\,\Delta\tilde{z}_j = \tilde{r}_{2,i}, \quad i = 1, 2, \ldots, p \tag{20b}$$

where $\tilde{\mathcal{J}}_i$ are small subsets of $\{1, 2, \ldots, p\}$.

Equivalent unconstrained problem

$$\underset{\Delta\tilde{z}}{\text{minimize}} \sum_{i=1}^{p} F_i(\Delta\tilde{z})$$

where

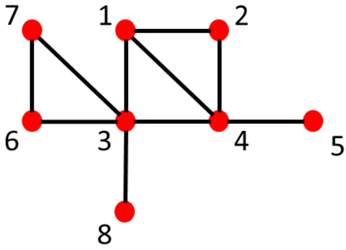
$$F_i(\Delta\tilde{z}) = \frac{1}{2}\Delta\tilde{z}_i^T \tilde{H}_i \Delta\tilde{z}_i - \tilde{r}_{1,i}^T \Delta\tilde{z}_i + I_{\mathcal{D}_i}(\Delta\tilde{z})$$

with $I_{\mathcal{D}_i}$ the indicator function for the set described by the ith equality constraint.

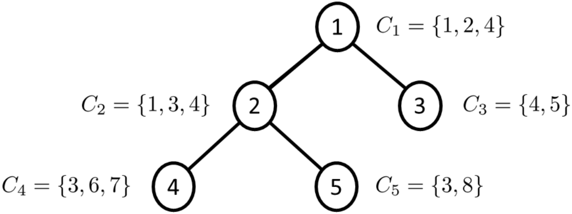
Simple Example

$$\begin{aligned}\underset{x}{\text{minimize}} \quad & \bar{F}_1(x_1, x_3) + \bar{F}_2(x_1, x_2, x_4) + \\ & \bar{F}_3(x_4, x_5) + \bar{F}_4(x_3, x_4) + \bar{F}_5(x_3, x_6, x_7) + \bar{F}_6(x_3, x_8). \quad (21)\end{aligned}$$

Has sparsity graph (edge between vertices if components in same term)



Clique Tree for Sparsity Graph



We now assign one computational agent for each clique, and we may assign $\bar{F}_i$ to an agent if and only if the indeces of its variables belong to the corresponding clique. Hence we can assign $\bar{F}_1 + \bar{F}_4$ to C_2, $\bar{F}_2$ to C_1, $\bar{F}_3$ to C_3, $\bar{F}_5$ to C_4 and $\bar{F}_6$ to C_5. (Not unique assignment)

Message Passing or Dynamic Programming over Trees

Start with the leaves and compute for agents 3, 4, and 5

$$m_{31}(x_4) = \min_{x_5} \left\{ \bar{F}_3(x_4, x_5) \right\} \tag{22}$$

$$m_{42}(x_3) = \min_{x_6, x_7} \left\{ \bar{F}_5(x_3, x_6, x_7) \right\} \tag{23}$$

$$m_{52}(x_3) = \min_{x_8} \left\{ \bar{F}_6(x_3, x_8) \right\} \tag{24}$$

Then add the results from agents 4 and 5 to the functions of Agent 2 and compute

$$m_{21}(x_1, x_4) = \min_{x_3} \left\{ \bar{F}_1(x_1, x_3) + \bar{F}_4(x_3, x_4) + m_{42}(x_3) + m_{52}(x_3) \right\} \tag{25}$$

Finally add the results from agents 2 and 3 to the functions of Agent 1 and compute

$$\min_{x_1, x_2, x_4} \left\{ \bar{F}_2(x_1, x_2, x_4) + m_{31}(x_4) + m_{21}(x_1, x_4) \right\}$$

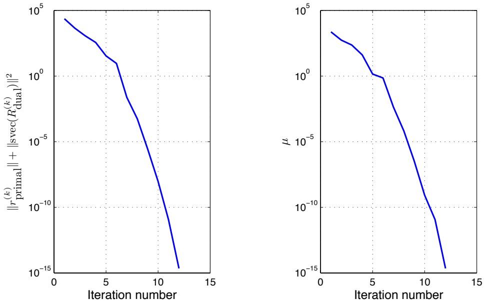
Comments

- Not easy in general to compute messages or value functions $m_{i,j}$.
- For linearly constrained convex quadratic problems the messages are convex quadratic functions with equality constraints.
- The dual variables can also be recovered.
- In fact results in a *multi-frontal factorization technique* for the KKT saddle point problem.

Comments ctd.

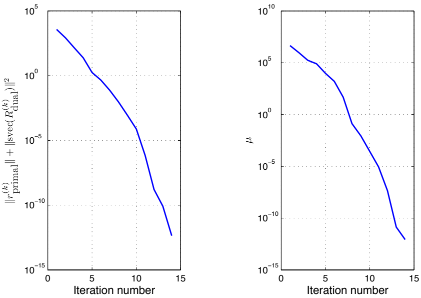
- The cliques for the search directions of the dual problem obtained using domain-space decomposition will not be the same as the cliques in the domain-space decomposition itself.
- They can however be obtained by “merging” cliques, where one clique might have to be merged with several others.
- All other computations in an IP algorithm also distribute over the clique tree.
- In total 6 upward and 6 downward passes through the clique tree, of which only one pass involves significant computations, for each iteration in an IP algorithm

Chain of 100 Uncertain Systems



- 198 cliques
- Height of qlique tree 99
- Largest clique of dimension 8.
- Each agent computed a factorization 12 times and needed to communicate with its neighbours 144 times.
- Dimension of marix to factorize was at most 62.
- Each agent had at most 2 neighbours.

The Scale-Free Network



- 579 cliques
- Height of qlique tree 35
- Largest clique of dimension 162.
- Each agent computed a factorization 14 times and needed to communicate with its neighbours 168 times.
- Dimension of marix to factorize was at most 5456.
- Each agent had at most 39 neighbours.

Summary

- Presented *scalable distributed* optimization algorithms that respect *privacy*.
- However, distributed solutions more costly when implemented centralized and especially so for second order methods.
- Robustness analysis has applications in power grids.
- Distributed localization of scattered sensor networks.
- Distributed predictiv control of platoons of vehicles.
- Distributed inertial motion capture

Acknowledgements

- Based on the thesis work by *Sina Khoshfetrat Pakazad* (LIU)
- Collaboration with Martin Andersen (DTU)and Anders Rantzer (LU)

Publications

S. Khoshfetrat Pakazad, “Divide and Conquer: Distributed Optimization and Robustness Analysis”, Linköping Studies in Science and Technology, Dissertations, No 1676, 2015.

S. Khoshfetrat Pakazad, M. S. Andersen, and A. Hansson. “Distributed soloutions for loosely coupled feasibility problems using proximal splitting methods.”, *Optimization Methods and Software*, 30(1):128–161, 2015.

S. Khoshfetrat Pakazad, A. Hansson, M. S. Andersen, and I. Nielsen. “Distributed primal-dual interior-point methods for solving tree-structured coupled convex problems using message-passing”, *Optimization Methods and Software*, DOI:10.1080/10556788.2016.1213839, 2016

M. S. Andersen, S. Khoshfetrat Pakazad, A. Hansson, and A. Rantzer, “Robust stability analysis of sparsely interconnected uncertain systems”, *IEEE Transactions on Automatic Control*, 59(8):2151–2156, 2014.

M, Kok, S. Khoshfetrat Pakazad, Th. B. Schön, A. Hansson, J. D. Hol, “A Scalable and Distributed Solution to the Inertial Motion Capture Problem”, arXiv:1603.06443, presented at Fusion, 2016.

S. Khoshfetrat Pakazad, E. Özkan, C. Fritsche, A. Hansson, and F. Gustafsson, “Distributed Localization of Tree-Structured Scattered Sensor Networks”, arXiv:1607.04798, 2016

Decomposition and Product Space Formulation

Feasibility problem from range-space decomposition can with $v = (x, U)$ be phrased as

$$\text{find} \quad v \tag{26a}$$

$$\text{subject to} \quad v \in \mathcal{C}_j, \quad j = 1, \ldots, p, \tag{26b}$$

where

$$\mathcal{C}_j = \left\{ v \middle| F_j^0 + \sum_{i=1}^{m} x_i F_j^i + G_j(U) \succeq 0 \right\}$$

Let

$$\bar{\mathcal{C}}_j = \{ s^j \in \mathbb{R}^{|\mathcal{J}_j|} \mid E_{\mathcal{J}_j}^T s^j \in \mathcal{C}_i \}, \quad j = 1, \ldots, p, \tag{27}$$

such that $s^j \in \bar{\mathcal{C}}_j$ implies $E_{\mathcal{J}_j}^T s^j \in \mathcal{C}_j$, where $E_{\mathcal{J}_j}$ are composed of rows of the identity matrix indexed by the set $\mathcal{J}_j$, which is the set of i such that v_i is constrained by $\mathcal{C}_j$. Let $\mathcal{I}_i = \{k | i \in \mathcal{J}_k\}$, i.e. the set of indeces of constraints, which depends on v_i.

Example revisited

Find $(x, u) = (x_1, x_2, x_3, x_4, u)$ such that

$$\begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 + u \end{bmatrix} \succeq 0 \ \& \ \begin{bmatrix} -u & x_3 \\ x_3 & x_4 \end{bmatrix} \succeq 0$$

Hence

$$\mathcal{J}_1 = \{1, 2, 5\}; \ \mathcal{J}_2 = \{3, 4, 5\}$$

and

$$\mathcal{I}_1 = \{1\}; \ \mathcal{I}_2 = \{1\}; \ \mathcal{I}_3 = \{2\}; \ \mathcal{I}_4 = \{2\}; \ \mathcal{I}_5 = \{1, 2\}$$

Product Space Formulation

Then (26) is equivalent to

$$\begin{aligned} \text{find} \quad & s^1, s^2, \ldots, s^p, v && (28a) \\ \text{subject to} \quad & s^i \in \bar{\mathcal{C}}_i, \quad i = 1, \ldots, p && (28b) \\ & s^i = E_{\mathcal{J}_i} v, \quad i = 1, \ldots, p && (28c) \end{aligned}$$

or

$$\begin{aligned} \text{find} \quad & S \\ \text{subject to} \quad & S \in \mathcal{C}, \; S \in \mathcal{D} \end{aligned} \qquad (29)$$

where

$$\begin{aligned} S &= (s^1, \ldots, s^p) \in \mathbb{R}^{|\mathcal{J}_1|} \times \cdots \times \mathbb{R}^{|\mathcal{J}_p|} \\ \mathcal{C} &= \bar{\mathcal{C}}_1 \times \cdots \times \bar{\mathcal{C}}_p \\ \mathcal{D} &= \{\bar{E} v \mid v \in \mathbb{R}^n\} \\ \bar{E} &= \begin{bmatrix} E_{\mathcal{J}_1}^T & \cdots & E_{\mathcal{J}_p}^T \end{bmatrix}^T. \end{aligned}$$

Example revisited

$$
\begin{aligned}
\text{find} \quad & s^1, s^2, v && \text{(30a)}\\
\text{subject to} \quad & s^1 \in \left\{ s^1 \middle| \begin{bmatrix} s_1^1 & s_2^1 \\ s_2^1 & s_1^1 + s_3^1 \end{bmatrix} \succeq 0 \right\}, && \text{(30b)}\\
& s^2 \in \left\{ s^2 \middle| \begin{bmatrix} -s_3^2 & s_1^2 \\ s_1^2 & s_2^2 \end{bmatrix} \succeq 0 \right\}, && \text{(30c)}\\
& s^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} v && \text{(30d)}\\
& s^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} v && \text{(30e)}
\end{aligned}
$$

Convex Minimization Formulation

Consider

$$\underset{S}{\text{minimize}} \quad F(S) := \frac{1}{2}\|S - P_{\mathcal{C}}(S)\|^2 + \frac{1}{2}\|S - P_{\mathcal{D}}(S)\|^2, \qquad (31)$$

where $P_{\mathcal{C}}(S)$ is the projection of S on the set $\mathcal{C}$ and similarly for $\mathcal{D}$.

This problem provides a solution to (29) if the optimal value is zero. A non-zero optimal value proves that (29) is infeasible.

Splitting

We equivalently write the problem with $x = S$ (new meaning of x) as

$$\underset{x,y}{\text{minimize}}\ f_1(x) + f_2(y) \tag{32a}$$

$$\text{subject to } x = y \tag{32b}$$

where

$$f_1(x) = \frac{1}{2}\|x - P_{\mathcal{C}}(x)\|^2; \quad f_2(x) = \frac{1}{2}\|x - P_{\mathcal{D}}(x)\|^2$$

Proximity Operator

Given a closed convex function $f : \mathbb{R}^n \to \mathbb{R}$, then for every $x \in \mathbb{R}^n$ the proximity operator of the function f, $\mathrm{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$ is defined as the unique minimizer of the following optimization problem,

$$\underset{y}{\text{minimize}} \quad f(y) + \frac{1}{2}\|x - y\|^2.$$

Alternating Linearization Methods

Algorithm 1 ALM

Given $y^{(1)}$
for $k = 1, 2 \ldots$ **do**
  $x^{(k+1)} = \text{prox}_{f_1}(y^{(k)} - \nabla f_2(y^{(k)}))$
  $y^{(k+1)} = \text{prox}_{f_2}(x^{(k+1)} - \nabla f_1(x^{(k+1)}))$
end for

where

$$\text{prox}_{f_1}(x) = \frac{x + P_{\mathcal{C}}(x)}{2}; \quad \text{prox}_{f_2}(x) = \frac{x + P_{\mathcal{D}}(x)}{2}$$

and

$$\nabla f_1(x) = x - P_{\mathcal{C}}(x); \quad \nabla f_2(x) = x - P_{\mathcal{D}}(x)$$

Distributed Implementation

Since

$$(P_{\mathcal{C}}(x))_i = P_{\bar{\mathcal{C}}_i}(x^i)$$

these projections can be distributed over p computatational agents.

Moreover

$$P_{\mathcal{D}}(x) = \bar{E}\left(\bar{E}^T\bar{E}\right)^{-1}\bar{E}^T x$$

where $\bar{E}^T\bar{E} = \mathrm{diag}(|\mathcal{I}_i|)$. Thus for the example

$$(P_{\mathcal{D}}(x))_1 = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1/2 \end{bmatrix} x^1 + \begin{bmatrix} & & \\ & & \\ & & 1/2 \end{bmatrix} x^2$$

$$(P_{\mathcal{D}}(x))_2 = \begin{bmatrix} & & \\ & & \\ & & 1/2 \end{bmatrix} x^1 + \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1/2 \end{bmatrix} x^2$$

and hence this projection requires information from neighboring computational agents which have variables in common.