Exploiting Chordality in Optimization Algorithms for Model Predictive Control

Anders Hansson

Division of Automatic Control
Linköping University

March 8, 2018

Outline

Dynamic Programming over Trees

Interior-Point Methods

Parametric QPs

Model Predictive Control

Parallel MPC

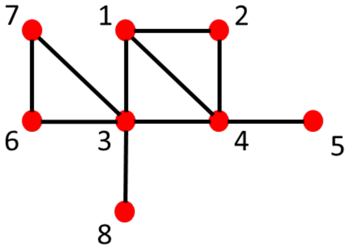
Stochastic MPC

Summary

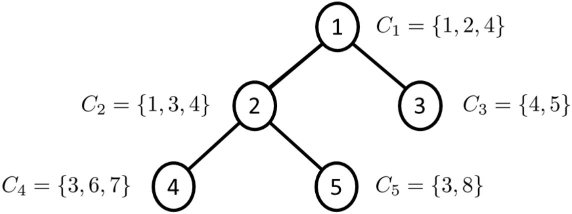
Simple Example

$$\underset{x}{\text{minimize}} \quad \bar{F}_1(x_1, x_3) + \bar{F}_2(x_1, x_2, x_4) + \bar{F}_3(x_4, x_5) + \bar{F}_4(x_3, x_4) + \bar{F}_5(x_3, x_6, x_7) + \bar{F}_6(x_3, x_8). \quad (1)$$

Has sparsity graph (edge between vertexes 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 indexes 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

$$m_{31}(x_4) = \min_{x_5} \left\{ \bar{F}_3(x_4, x_5) \right\} \quad (2)$$

$$m_{42}(x_3) = \min_{x_6, x_7} \left\{ \bar{F}_5(x_3, x_6, x_7) \right\} \quad (3)$$

$$m_{52}(x_3) = \min_{x_8} \left\{ \bar{F}_6(x_3, x_8) \right\} \quad (4)$$

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

$$m_{21}(x_1, x_4) = \min_{x_3} \left\{ \bar{F}_1(x_1, x_3) + \bar{F}_4(x_3, x_4) + m_{42}(x_3) + m_{52}(x_3) \right\} \quad (5)$$

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.
- Can be used to compute search directions in IP methods
- 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

Interior-Point Methods

Consider the QP

$$\min_z \frac{1}{2} z^T \mathcal{Q} z + q^T z \tag{6}$$

$$\text{s.t. } \mathcal{A} z = b \tag{7}$$

$$\mathcal{D} z \leq e \tag{8}$$

where $\mathcal{Q} \succeq 0$, and $\mathcal{A}$ has full row rank.

KKT optimality conditions:

$$\begin{bmatrix} \mathcal{Q} & \mathcal{A}^T & \mathcal{D}^T & \\ \mathcal{A} & & & \\ \mathcal{D} & & & I \\ & & & M \end{bmatrix} \begin{bmatrix} z \\ \lambda \\ \mu \\ s \end{bmatrix} = \begin{bmatrix} -q \\ b \\ e \\ 0 \end{bmatrix} \tag{9}$$

and $(\mu, s) \geq 0$, where $M = \text{diag}(\mu)$.

Search Directions

Linearize:

$$\begin{bmatrix} \mathcal{Q} & \mathcal{A}^T & \mathcal{D}^T & \\ \mathcal{A} & & & \\ \mathcal{D} & & & I \\ & & S & M \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta \lambda \\ \Delta \mu \\ \Delta s \end{bmatrix} = \begin{bmatrix} r_z \\ r_\lambda \\ r_\mu \\ r_s \end{bmatrix} \tag{10}$$

where $S = \text{diag}(s)$, and where $r = (r_z, r_\lambda, r_\mu, r_s)$ is residual vector.

Reduced KKT system

Equivalently $\Delta s = r_\mu - \mathcal{D}\Delta z$, $\Delta\mu = S^{-1}(r_s - M\Delta s)$ and

$$\begin{bmatrix} \mathcal{Q} + \mathcal{D}^T S^{-1} M \mathcal{D} & \mathcal{A}^T \\ \mathcal{A} & \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} r_z - \mathcal{D}^T S^{-1}(r_s - M r_\mu) \\ r_\lambda \end{bmatrix}. \tag{11}$$

Unique solution iff

$$\mathcal{Q}_s = \mathcal{Q} + \mathcal{D}^T S^{-1} M \mathcal{D} \tag{12}$$

is positive definite on the null-space of $\mathcal{A}$.

Parametric QPs

Consider

$$\min_z \frac{1}{2} z^T M z + m^T z \tag{13}$$

$$\text{s.t. } Cz = d \tag{14}$$

with C full row rank and $M \succeq 0$.

KKT conditions:

$$\begin{bmatrix} M & C^T \\ C & \end{bmatrix} \begin{bmatrix} z \\ \lambda \end{bmatrix} = \begin{bmatrix} -m \\ d \end{bmatrix}.$$

with unique solution if and only if $M + C^T C \succ 0$.

Partitioned Problem

Let

$$M = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}; \quad C = \begin{bmatrix} A & B \\ & D \end{bmatrix}; \quad d = \begin{bmatrix} e \\ f \end{bmatrix}; \quad m = \begin{bmatrix} q \\ r \end{bmatrix}; \quad z = \begin{bmatrix} x \\ y \end{bmatrix}$$

with A full row rank.

Solve

$$\min_x \frac{1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + q^T x \tag{15}$$

$$\text{s.t. } Ax + By = e \tag{16}$$

parametrically with respect to all y.

KKT Conditions for Parametric Problem

$$\begin{bmatrix} Q & A^T \\ A & \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} -q - Sy \\ e - By \end{bmatrix}.$$

- Solution x will be affine in y
- Results in a quadratic message in y.
- The 1,1-block of $M + C^T C$ is $Q + A^T A$, which by the Schur complement formula is positive definite, which implies unique solution

Rank Condition

In case A does not have full row rank, perform a rank-revealing factorization

$$\begin{bmatrix} \bar{A}_1 \\ 0 \end{bmatrix} x + \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix} y = \begin{bmatrix} \bar{e}_1 \\ \bar{e}_2 \end{bmatrix}$$

and append the constraint $\bar{B}_2 y = \bar{e}_2$ to belong to

$$Dy = f.$$

Model Predictive Control (MPC)

$$\min_u \frac{1}{2} \sum_{k=0}^{N-1} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T Q \begin{bmatrix} x_k \\ u_k \end{bmatrix} + \frac{1}{2} x_N^T S x_N \tag{17}$$

$$\text{s.t. } x_{k+1} = Ax_k + Bu_k, \quad x_0 = \bar{x} \tag{18}$$

where $Q \succeq 0$ and $S \succeq 0$

Let $\mathcal{I}_{\mathcal{C}_k}(x_k, u_k, x_{k+1})$ be indicator function for

$$\mathcal{C}_k = \{(x_k, u_k, x_{k+1}) \mid x_{k+1} = Ax_k + Bu_k\}$$

and $\mathcal{I}_{\mathcal{D}}(x_0)$ indicator function for

$$\mathcal{D} = \{x_0 \mid x_0 = \bar{x}\}$$

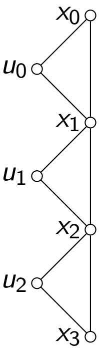
Equivalent Formulation

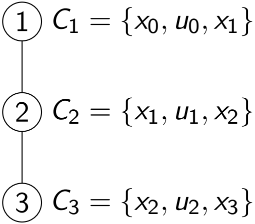
$$\min_{x} \quad \bar{F}_1(x_0, u_0, x_1) + \cdots + \bar{F}_N(x_{N-1}, u_{N-1}, x_N), \tag{19}$$

where

$$\begin{aligned}
\bar{F}_1(x_0, u_0, x_1) &= \mathcal{I}_{\mathcal{D}}(x_0) + \frac{1}{2}\begin{bmatrix} x_0 \\ u_0 \end{bmatrix}^T Q \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} + \mathcal{I}_{\mathcal{C}_0}(x_0, u_0, x_1) \\
\bar{F}_{k+1}(x_k, u_k, x_{k+1}) &= \frac{1}{2}\begin{bmatrix} x_k \\ u_k \end{bmatrix}^T Q \begin{bmatrix} x_k \\ u_k \end{bmatrix} + \mathcal{I}_{\mathcal{C}_k}(x_k, u_k, x_{k+1}) \\
\bar{F}_N(x_{N-1}, u_{N-1}, x_N) &= \frac{1}{2}\begin{bmatrix} x_{N-1} \\ u_{N-1} \end{bmatrix}^T Q \begin{bmatrix} x_{N-1} \\ u_{N-1} \end{bmatrix} + \mathcal{I}_{\mathcal{C}_{N-1}}(x_{N-1}, u_{N-1}, x_N) \\
&\quad + \frac{1}{2} x_N^T S x_N
\end{aligned}$$

Sparsity Graph and Clique Tree





Assign $\bar{F}_k$ to C_k.

Can just as well take C_2 or C_3 as root!

Parallel Computations

Same problem as before but with $N = 6$.

Dummy variables $\bar{u}_0$ and $\bar{u}_1$ and consensus constraints:

$$\bar{u}_0 = x_3, \quad \bar{u}_1 = x_6.$$

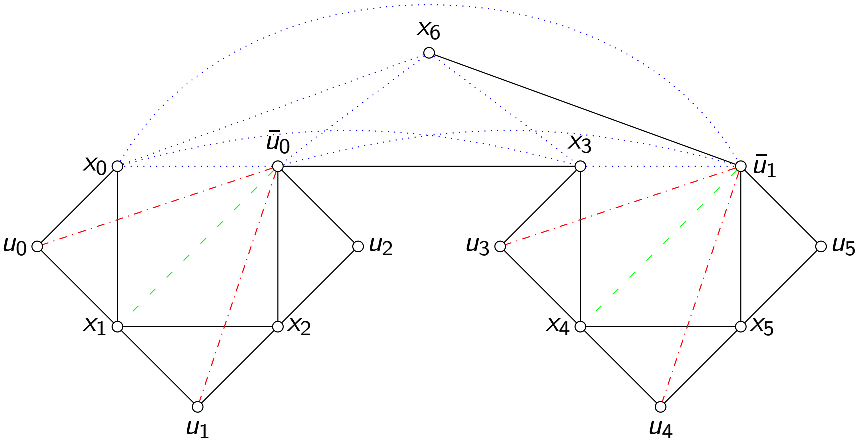
Similar to Nielsen (2017). Define

$$\begin{aligned}
\mathcal{C}_{-1} &= \{x_0 : x_0 = \bar{x}\} \\
\mathcal{C}_k &= \{(x_k, u_k, x_{k+1}) : \ x_{k+1} = Ax_k + Bu_k\}; \ k = 0, 1 \\
\mathcal{C}_2 &= \{(x_2, u_2, \bar{u}_0) : \ \bar{u}_0 = Ax_2 + Bu_2\} \\
\mathcal{C}_k &= \{(x_k, u_k, x_{k+1}) : \ x_{k+1} = Ax_k + Bu_k\}; \ k = 3, 4 \\
\mathcal{C}_5 &= \{(x_5, u_5, \bar{u}_1) : \ \bar{u}_1 = Ax_5 + Bu_5\} \\
\mathcal{D}_0 &= \{(x_3, \bar{u}_0) : \ \bar{u}_0 = x_3\} \\
\mathcal{D}_1 &= \{(x_6, \bar{u}_1) : \ \bar{u}_1 = x_6\}.
\end{aligned} \tag{20}$$

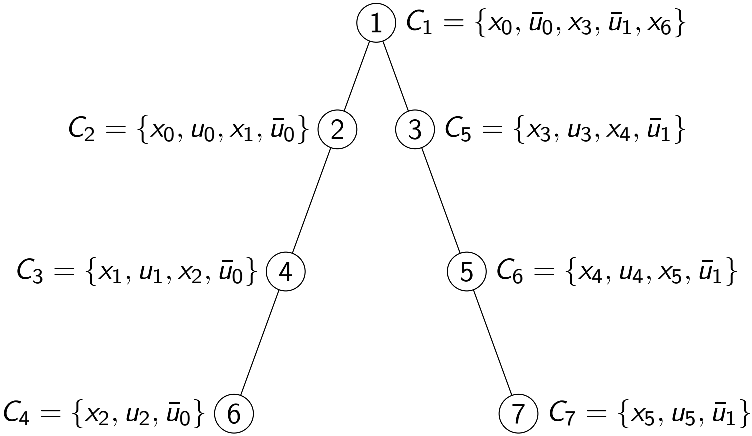
Equivalent Problem

$$
\begin{aligned}
\min_u \ & \frac{1}{2}\sum_{k=0}^{1}\begin{bmatrix}x_k\\u_k\end{bmatrix}^T Q \begin{bmatrix}x_k\\u_k\end{bmatrix} + \mathcal{I}_{\mathcal{C}_k}\{x_k, u_k, x_{k+1}\} + \\
& \frac{1}{2}\begin{bmatrix}x_2\\u_2\end{bmatrix}^T Q \begin{bmatrix}x_2\\u_2\end{bmatrix} + \mathcal{I}_{\mathcal{C}_2}\{x_2, u_2, \bar{u}_0\} + \\
& \frac{1}{2}\sum_{k=3}^{4}\begin{bmatrix}x_k\\u_k\end{bmatrix}^T Q \begin{bmatrix}x_k\\u_k\end{bmatrix} + \mathcal{I}_{\mathcal{C}_k}\{x_k, u_k, x_{k+1}\} + \\
& \frac{1}{2}\begin{bmatrix}x_5\\u_5\end{bmatrix}^T Q \begin{bmatrix}x_5\\u_5\end{bmatrix} + \mathcal{I}_{\mathcal{C}_5}\{x_5, u_5, \bar{u}_1\} + \frac{1}{2}\bar{u}_1^T S \bar{u}_1 + \\
& \mathcal{I}_{\mathcal{C}_{-1}}\{x_0\} + \mathcal{I}_{\mathcal{D}_0}\{x_3, \bar{u}_0\} + \mathcal{I}_{\mathcal{D}_1}\{x_6, \bar{u}_1\}
\end{aligned} \tag{21}
$$

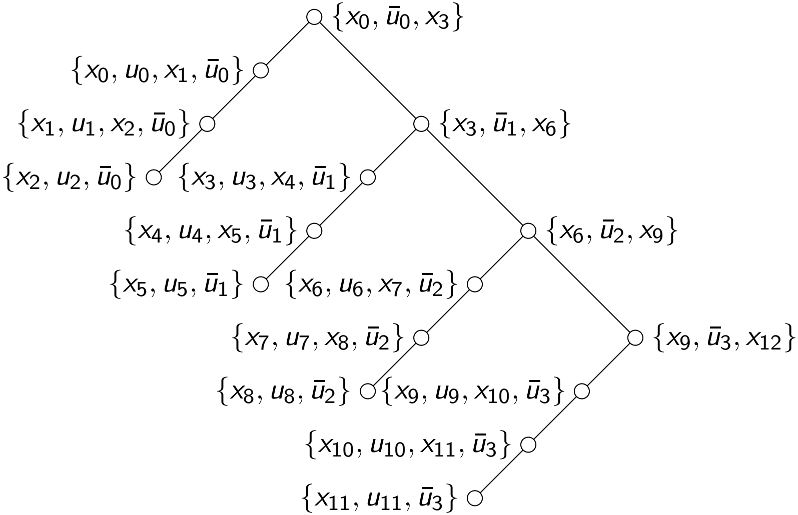
Sparsity Graph



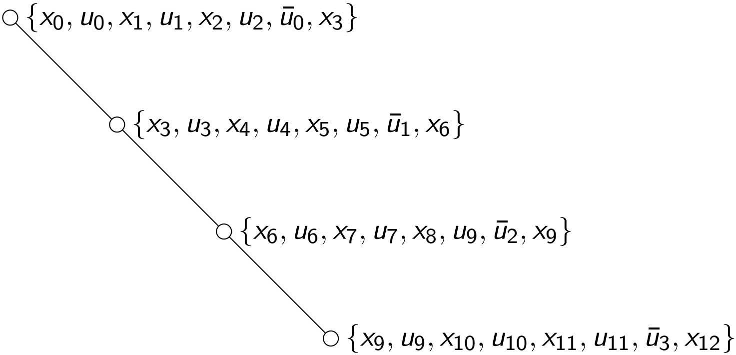
Clique Tree


1
$C_1 = \{x_0, \bar{u}_0, x_3, \bar{u}_1, x_6\}$
$C_2 = \{x_0, u_0, x_1, \bar{u}_0\}$
2
3
$C_5 = \{x_3, u_3, x_4, \bar{u}_1\}$
$C_3 = \{x_1, u_1, x_2, \bar{u}_0\}$
4
5
$C_6 = \{x_4, u_4, x_5, \bar{u}_1\}$
$C_4 = \{x_2, u_2, \bar{u}_0\}$
6
7
$C_7 = \{x_5, u_5, \bar{u}_1\}$

Clique Tree with Four Parallel Branches



Merging of Cliques



Stochastic MPC

- d is the number of stochastic events that can take place at each time stage k
- r is the number of time stages for which we consider stochastic events to take place.
- Outcomes of the stochastic events are the different values of A_k^j, B_k^j and v_k^j.
- Number of scenarios is $M = d^r$

Optimization Problem

$$\min_{u} \sum_{j=1}^{M} \omega_j \left(\frac{1}{2} \sum_{k=0}^{N-1} \begin{bmatrix} x_k^j \\ u_k^j \end{bmatrix}^T Q \begin{bmatrix} x_k^j \\ u_k^j \end{bmatrix} + \frac{1}{2} (x_N^j)^T S x_N^j \right) \tag{22}$$

$$\text{s.t. } x_{k+1}^j = A_k^j x_k^j + B_k^j u_k^j + v_k^j, \quad x_0^j = \bar{x} \tag{23}$$

$$\bar{C} u = 0 \tag{24}$$

where $u = (u^1, u^2, \dots, u^M)$ with $u^j = (u_0^j, u_1^j, \dots, u_{N-1}^j)$, and

$$\bar{C} = \begin{bmatrix} C_{1,2} & -C_{1,2} & & & \\ & C_{2,3} & -C_{2,3} & & \\ & & \ddots & \ddots & \\ & & & C_{M-1,M} & -C_{M-1,M} \end{bmatrix}$$

with

$$C_{j,j+1} = \begin{bmatrix} I & 0 \end{bmatrix}$$

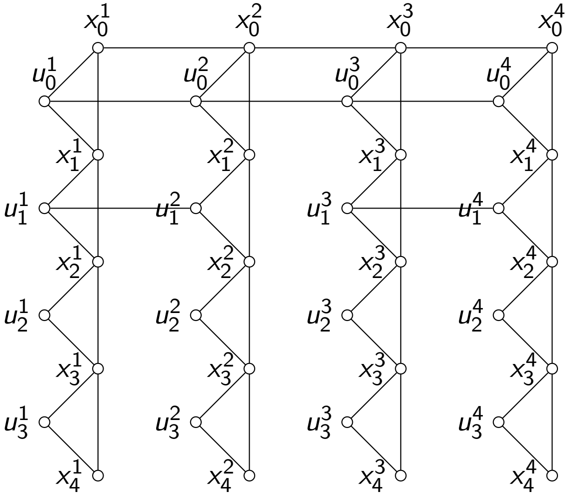
The constraint $\bar{C} u = 0$ is the *non-ancipativity constraint*.

Comments

- Several of A_k^j, B_k^j and v_k^j are the same
- ω_j is the probability of scenario j
- Instead of $x_0^j = \bar{x}$ we equivalently write $x_0^1 = \bar{x}$ and $x_0^j = x_0^{j+1}$, for $1 \leq j \leq M - 1$.

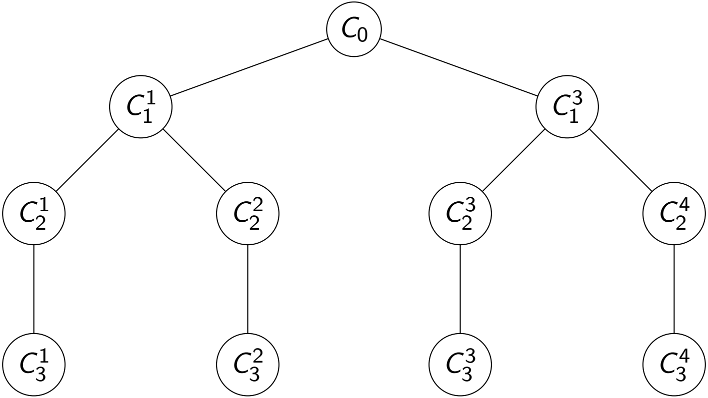
Sparsity Graph

Case of $N = 4,\ d = r = 2 \Rightarrow M = 4$



Make chordal embedding.

Clique Tree



Summary

- Interior-point methods over trees based on dynamic programming or message passing to compute search directions.
- Needs less communication than other distributed algorithms
- More complicated than first order methods
- Model predictive control (MPC)
- Parallel MPC
- Stochastic MPC
- Also regularized MPC and robust MPC possible.

Acknowledgements

Collaboration with *Sina Khoshfetrat Pakazad*

Publications

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

A. Hansson and S. Khoshfetrat Pakazad. "Exploiting Chordality in Optimization Algorithms for Model Predictive Control", arXiv:1711.10254, 2017

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

S. Khoshfetrat Pakazad, A. Hansson, M. S. Andersen, a. Rantzer. "Distributed Semidefinite Programming with Application to Large-scale System Analysis", IEEE Transactions on Automatic Control, DOI: 10.1109/TAC.2017.2739644, 2017.

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, "Divide and Conquer: Distributed Optimization and Robustness Analysis", Linköping Studies in Science and Technology, Dissertations, No 1676, 2015