ADMM for General Convex QPs - Optimal Convergence, Infeasibility Detection and Acceleration

Arvind U. Raghunathan   Stefano Di Cairano
Mitsubishi Electric Research Laboratories

EMBOPT 2014
8 September, 2014

Convex Quadratic Program

$$\begin{aligned} \min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned} \tag{QP}$$

- $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{Q} \succeq 0$ – symmetric positive semidefinite
- $\mathcal{Y}$ – closed, convex set with **simple** projection operation
  - Box constraints $\mathcal{Y} = [\mathbf{y}^\mathbf{l}, \mathbf{y}^\mathbf{u}]$
  - General inequalities $\mathcal{Y} = \{\mathbf{y} | \mathbf{B}\mathbf{y} \leq \mathbf{c}\}$
    - $\mathbf{B}\mathbf{y} + \mathbf{z} = \mathbf{c}$, $\mathcal{Y}' = \mathbb{R}^N \times \mathbb{R}^m_+$
    - $\mathbf{y}' = (\mathbf{y}, \mathbf{z})$, $\mathbf{A}\mathbf{y}' = (\mathbf{B}\ \mathbf{I}_m)\mathbf{y}' = \mathbf{c}$, $\mathbf{y}' \in \mathcal{Y}'$

Convex Quadratic Program

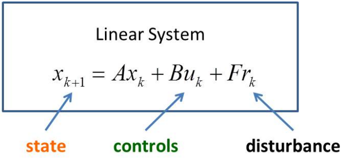
$$
\begin{aligned}
\min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\
\text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\
& \mathbf{y} \in \mathcal{Y}
\end{aligned}
\tag{QP}
$$

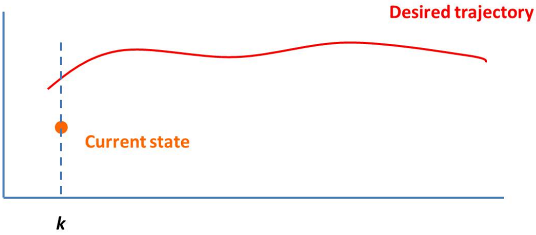
Assumptions

- $\mathbf{A}$ full row rank
- $\mathbf{Z}^T\mathbf{Q}\mathbf{Z} \succ 0$ (reduced Hessian) where $\mathbf{Z} = \text{null}(\mathbf{A})$

Our Focus – Model Predictive Control

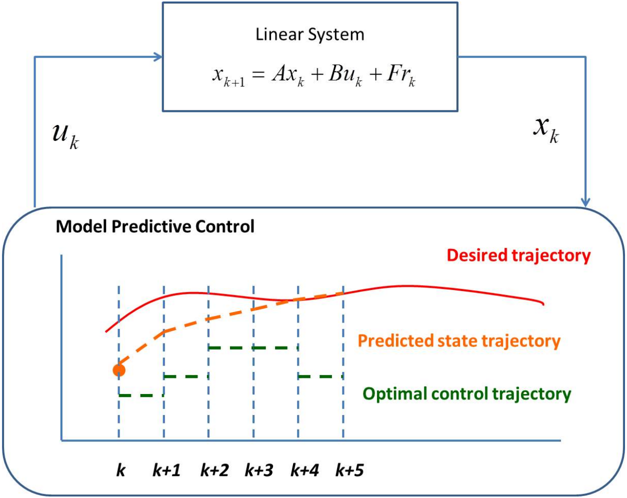
Numerous applications for QPs
– compressed sensing, control, estimation, engineering, finance, ...





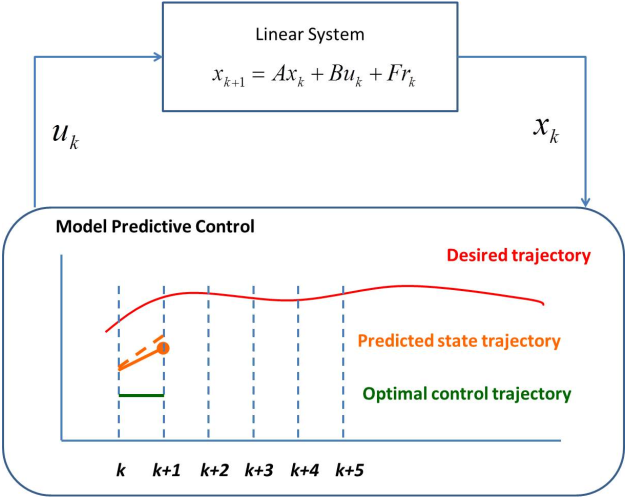
Our Focus - Model Predictive Control

Numerous applications for QPs
- compressed sensing, control, estimation, engineering, finance, ...



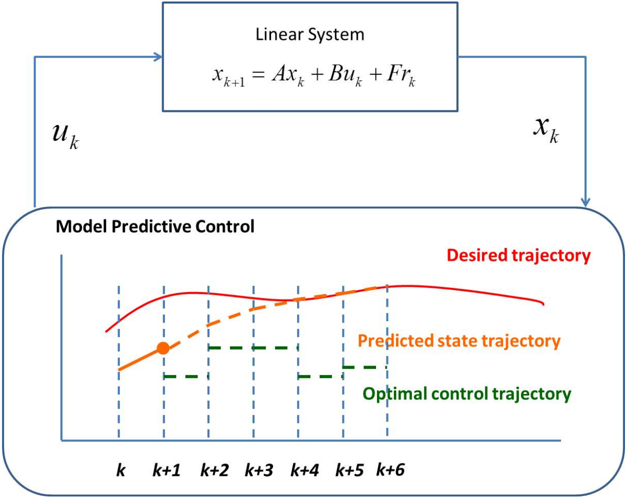
Our Focus – Model Predictive Control

Numerous applications for QPs
– compressed sensing, control, estimation, engineering, finance, ...



Our Focus - Model Predictive Control

Numerous applications for QPs
- compressed sensing, control, estimation, engineering, finance, ...



MPC Formulation

$$
\begin{aligned}
\min_{\{x_t\}_{t=0}^{N},\{u_t\}_{t=0}^{N-1}} \; & \frac{1}{2}\sum_{t=0}^{N-1}\left(x_t^T Q_x x_t + u_t^T R u_t\right) + \frac{1}{2} x_N^T P x_N \\
\text{s.t. } & x_{t+1} = A x_t + B u_t + F r_t \\
& (x_{t+1}, u_t) \in \mathcal{X} \times \mathcal{U} \\
& r_t = r(k+t) \\
& x_0 = x(k)
\end{aligned}
$$

Typically, $R \succ 0$ positive definite.

MPC Formulation

$$\mathbf{y} = (x_1, \ldots, x_N, u_0, \ldots, u_{N-1})$$

$$\mathbf{Q} = \begin{pmatrix} \mathbf{I}_{N-1} \otimes Q_x & 0 & 0 \\ 0 & Q_N & 0 \\ 0 & 0 & \mathbf{I}_N \otimes R \end{pmatrix}$$

$$\mathbf{b} = (Ax(k) + Fd_0, \ldots, Fd_{N-1})$$

$$\mathcal{Y} = \underbrace{\mathcal{X} \times \ldots \mathcal{X}}_{N \text{ times}} \times \underbrace{\mathcal{U} \times \ldots \mathcal{U}}_{N \text{ times}}$$

$$\mathbf{A} = \left(\begin{array}{ccccc|c} \mathbf{I}_{n_x} & 0 & \ldots & 0 & 0 & \\ -A & \mathbf{I}_{n_x} & \ldots & 0 & 0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots & -\mathbf{I}_N \otimes B \\ 0 & 0 & \ldots & \mathbf{I}_{n_x} & 0 & \\ 0 & 0 & \ldots & -A & I_{n_x} & \end{array}\right)$$

MPC Formulation

$$\min_{\{x_t\}_{t=0}^{N},\{u_t\}_{t=0}^{N-1}} \frac{1}{2}\sum_{t=0}^{N-1}\left(x_t^T Q_x x_t + u_t^T R u_t\right) + \frac{1}{2}x_N^T P x_N$$
$$\text{s.t. } x_{t+1} = Ax_t + Bu_t + Fr_t$$
$$(x_{t+1}, u_t) \in \mathcal{X} \times \mathcal{U}$$
$$r_t = r(k+t)$$
$$x_0 = x(k)$$

Typically, $R \succ 0$ positive definite.
Eliminate x_t's using dynamics
$\implies$ Strictly Convex QP.
$\implies$ $\mathbf{Z}^T\mathbf{Q}\mathbf{Z} \succ 0$ (reduced hessian) where $\mathbf{Z} = \text{null}(\mathbf{A})$

MPC = Convex Quadratic Program

$$\begin{aligned} \min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned} \tag{QP}$$

Assumptions

- $\mathbf{A}$ full row rank
- $\mathbf{Z}^T\mathbf{Q}\mathbf{Z} \succ 0$ (reduced Hessian) where $\mathbf{Z} = \text{null}(\mathbf{A})$

MPC = Convex Quadratic Program

$$\begin{aligned} \min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned} \tag{QP}$$

Solve (QP) on Low Computing Power Platforms

- Matrix factorizations not possible
- Cannot run Interior Point Methods, Active-sets etc.
- Fast Gradient (Nesterov, Richter et al)
- Augmented Lagrangian (Necoara et al)
- Dual Fast Gradient (Patrinos et al)
- ADMM (Ghadimi et al, O'Donoghue et al, ...)

Remainder of the Talk

- ADMM algorithm for (QP)
  - β **is fixed** – like Ghadimi et al
  - $\mathcal{Y}$ is **always simple/computable projection**
- **Feasible & Infeasible** instances of (QP)
- Main convergence results
  - **2-step Q-Linear** convergence results
- Optimal parameter selection for ADMM
  - **Weaker** assumptions
  - $\mathcal{Y} = \{\mathbf{y} | \mathbf{B}\mathbf{y} \leq \mathbf{c}\}$ with $\mathbf{B}$ full row rank (Ghadimi et al)
  - **Excludes** finite lower and upper bounds
- Convergence Acceleration
  - **Conjugate Gradient** iterations

Outline

Augmented Lagrangian Formulation

$$\begin{aligned} \min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{w} \in \mathcal{Y} \\ & \mathbf{y} = \mathbf{w} \end{aligned}$$

- $\mathbf{y}$ – addresses Optimality
- $\mathbf{w}$ – addresses Feasibility
- Problem still coupled in $\mathbf{y}, \mathbf{w}$

Augmented Lagrangian Formulation

$$\begin{aligned} \min\ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} + \frac{\beta}{2}\|\mathbf{y}-\mathbf{w}-\boldsymbol{\lambda}\|^2 \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{w} \in \mathcal{Y} \end{aligned} \qquad \text{(AugQP)}$$

- $\beta > 0$ – parameter
- $\beta\boldsymbol{\lambda}$ – multipliers for convex set
- Problem still coupled in $\mathbf{y}, \mathbf{w}$
- ADMM – alternates b/w minimizing over $\mathbf{y}, \mathbf{w}$

ADMM Formulation

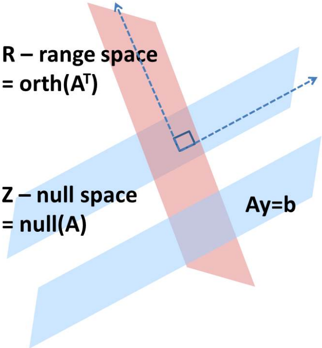
Solve for $\mathbf{y}^{k+1}$ - Optimality

$$\min_{\mathbf{y}} \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} + \frac{\beta}{2}\left\|\mathbf{y} - \mathbf{w}^k - \boldsymbol{\lambda}^k\right\|^2$$
$$\text{s.t. } \mathbf{A}\mathbf{y} = \mathbf{b}$$

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\boldsymbol{\xi}^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

For **fixed** β, need to factorize only **once**.

ADMM Formulation



Range & Null Space Decomposition

$$y = Ry_R + Zy_Z$$

Range Space Component

$$y_R = (AR)^{-1}b$$

Null Space Component

$$\min_{\mathbf{y_Z}} \left(\begin{array}{c} \frac{1}{2}(\mathbf{Zy_Z})^T \left(\mathbf{Q}/\beta + \mathbf{I}_n\right)(\mathbf{Zy_Z}) \\ + \left(\mathbf{q}/\beta - \mathbf{w}^k - \boldsymbol{\lambda}^k + \left(\mathbf{Q}/\beta + \mathbf{I}_n\right)\mathbf{Ry_R}\right)^T \mathbf{Zy_Z} \end{array} \right)$$

ADMM Formulation

Solve for $\mathbf{y}^{k+1}$ - Optimality

Solution:

$$\mathbf{y}^{k+1} = \mathbf{M}(\mathbf{w}^k + \boldsymbol{\lambda}^k) - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b}$$

where,

$$\mathbf{M} = \mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z}/\beta + \mathbf{I}_{n-m})^{-1}\mathbf{Z}, \implies 0 \preceq \mathbf{M} \prec \mathbf{I}_n$$

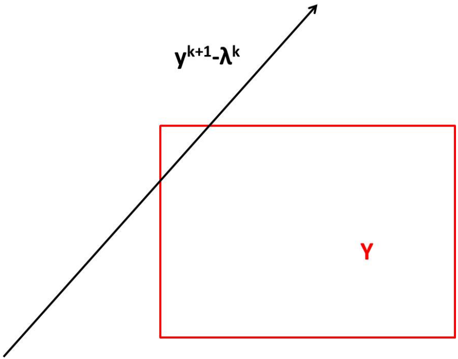
$$\mathbf{N} = (\mathbf{I}_n - \mathbf{M}\mathbf{Q}/\beta)\mathbf{R}(\mathbf{A}\mathbf{R})^{-1}\mathbf{b}$$

$\mathbf{M}, \mathbf{N}$ **pre-computed** for fixed β

ADMM Formulation

Solve for $\mathbf{w}^{k+1}$ – Set Feasibility

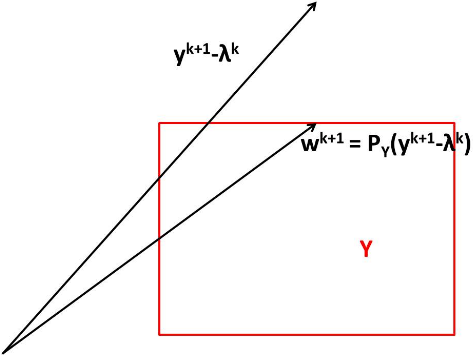
$$\min_{\mathbf{w}} \frac{\beta}{2} \left\| \mathbf{y}^{k+1} - \mathbf{w} - \boldsymbol{\lambda}^k \right\|^2$$
$$\text{s.t. } \mathbf{w} \in \mathcal{Y}$$



ADMM Formulation

Solve for $\mathbf{w}^{k+1}$ – Set Feasibility

$$\min_{\mathbf{w}} \frac{\beta}{2} \left\| \mathbf{y}^{k+1} - \mathbf{w} - \boldsymbol{\lambda}^k \right\|^2$$
$$\text{s.t. } \mathbf{w} \in \mathcal{Y}$$



ADMM Formulation

Solve for $\mathbf{y}^{k+1}$

$$\mathbf{y}^{k+1} = \mathbf{M}(\mathbf{w}^k + \boldsymbol{\lambda}^k) - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b}$$

Solve for $\mathbf{w}^{k+1}$

$$\mathbf{w}^{k+1} = \mathbb{P}_{\mathcal{Y}}(\mathbf{y}^{k+1} - \boldsymbol{\lambda}^k)$$

Update $\boldsymbol{\lambda}^{k+1}$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \mathbf{w}^{k+1} - \mathbf{y}^{k+1}$$

ADMM Algorithm

Choose $\mathbf{w}^0, \boldsymbol{\lambda}^0$. Set $k = 0$.

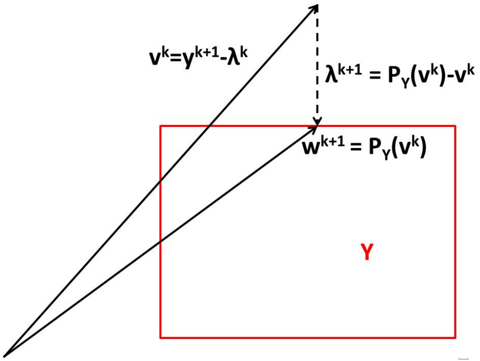
while (termination condition not satisfied) **do**

$$\begin{aligned} \mathbf{v}^k &= \mathbf{M}\mathbf{w}^k + (\mathbf{M} - \mathbf{I}_n)\boldsymbol{\lambda}^k - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b} \\ \mathbf{w}^{k+1} &= \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^k) \\ \boldsymbol{\lambda}^{k+1} &= \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^k) - \mathbf{v}^k \end{aligned}$$

Set $k = k + 1$

Variational Inequality Always Holds

$$\mathbf{w}^{k+1} = \arg\min_{\mathbf{w}\in\mathcal{Y}} \|\mathbf{w} - \mathbf{v}^k\|^2$$
$$\Longrightarrow (\mathbf{w}^{k+1} - \mathbf{v}^k)^T(\mathbf{w} - \mathbf{w}^{k+1}) \geq 0 \;\forall\, \mathbf{w} \in \mathcal{Y}$$
$$\Longrightarrow (\boldsymbol{\lambda}^{k+1})^T(\mathbf{w} - \mathbf{w}^{k+1}) \geq 0 \;\forall\, \mathbf{w} \in \mathcal{Y}$$
$$\Longrightarrow \mathbf{w}^{k+1} \in \mathcal{Y} \perp \boldsymbol{\lambda}^{k+1}$$



Outline

ADMM Formulation

Feasible Instances of (QP)

Infeasible Instances of (QP)

Convergence Acceleration

Optimality Conditions

Let $(\mathbf{y}^*, \mathbf{w}^*, \xi^*, \boldsymbol{\lambda}^*)$ be optimal solution of (QP).

$$
\begin{aligned}
\mathbf{Q}\mathbf{y}^* + \mathbf{A}^T\xi^* + \mathbf{q} - \boldsymbol{\lambda}^* &= 0 && \text{(Dual feasibility)}\\
\mathbf{A}\mathbf{y}^* &= \mathbf{b} && \text{(Subspace feasibility)}\\
\mathbf{w}^* \in \mathcal{Y} \perp &\ \boldsymbol{\lambda}^* && \text{(Variational Inequality)}\\
\mathbf{y}^* - \mathbf{w}^* &= 0 && \text{(Primal feasibility)}
\end{aligned}
$$

Optimality Conditions

Let $(\mathbf{y}^*, \mathbf{w}^*, \xi^*, \boldsymbol{\lambda}^*)$ be optimal solution of (QP).

$$
\begin{aligned}
\mathbf{Q}\mathbf{y}^* + \mathbf{A}^T\xi^* + \mathbf{q} - \boldsymbol{\lambda}^* &= 0 && \text{(Dual feasibility)} \\
\mathbf{A}\mathbf{y}^* &= \mathbf{b} && \text{(Subspace feasibility)} \\
\mathbf{w}^* \in \mathcal{Y} &\perp \boldsymbol{\lambda}^* && \text{(Variational Inequality)} \\
\mathbf{y}^* - \mathbf{w}^* &= 0 && \text{(Primal feasibility)}
\end{aligned}
$$

Subspace feasibility, Variational Inequality - Always Satisfied

$$\mathbf{A}\mathbf{y}^{k+1} = \mathbf{b}$$

$$\mathbf{w}^{k+1} \in \mathcal{Y} \perp \beta\boldsymbol{\lambda}^{k+1}$$

Termination Conditions

Primal Infeasibility

$$\|\mathbf{y}^{k+1} - \mathbf{w}^{k+1}\| = \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^{k}\|$$

Dual Infeasibility

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\boldsymbol{\xi}^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

$$\mathbf{Q}\mathbf{y}^{k+1} + \mathbf{A}^T\boldsymbol{\xi}^{k+1} + \mathbf{q} - \beta\boldsymbol{\lambda}^{k+1} = \beta\mathbf{w}^k + \beta\underbrace{\left(\boldsymbol{\lambda}^k - \mathbf{y}^{k+1} - \boldsymbol{\lambda}^{k+1}\right)}_{-\mathbf{w}^{k+1}}$$

$$\|\mathbf{Q}\mathbf{y}^{k+1} + \mathbf{A}^T\boldsymbol{\xi}^{k+1} + \mathbf{q} - \beta\boldsymbol{\lambda}^{k+1}\| = \beta\|\mathbf{w}^{k+1} - \mathbf{w}^k\|$$

Projection Operator-based Inequalities

Let $(\mathbf{y}^*, \mathbf{w}^*, \boldsymbol{\lambda}^*)$ be optimal solution of (QP).
Denote, $\mathbf{v}^* = \mathbf{M}\mathbf{w}^* + (\mathbf{M} - \mathbf{I}_n)\boldsymbol{\lambda}^* - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b}$

Inequality 1

$$\left\| \begin{pmatrix} \mathbf{w}^{k+1} \\ \boldsymbol{\lambda}^{k+1} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^* \\ \boldsymbol{\lambda}^* \end{pmatrix} \right\| = \left\| \begin{pmatrix} \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^k) \\ (\mathbb{P}_{\mathcal{Y}} - \mathbf{I}_n)(\mathbf{v}^k) \end{pmatrix} - \begin{pmatrix} \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^*) \\ (\mathbb{P}_{\mathcal{Y}} - \mathbf{I}_n)(\mathbf{v}^*) \end{pmatrix} \right\| \leq \|\mathbf{v}^k - \mathbf{v}^*\|$$

Inequality 2

$$\begin{aligned} \|\mathbf{v}^k - \mathbf{v}^*\| &= \left\| \mathbf{M}(\mathbf{w}^k - \mathbf{w}^*) + (\mathbf{M} - \mathbf{I}_n)(\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*) \right\| \\ &\leq \left\| \begin{pmatrix} \mathbf{w}^k \\ \boldsymbol{\lambda}^k \end{pmatrix} - \begin{pmatrix} \mathbf{w}^* \\ \boldsymbol{\lambda}^* \end{pmatrix} \right\| (\text{since } 0 \preceq \mathbf{M} \prec \mathbf{I}_n) \end{aligned}$$

ADMM - Convergence Analysis

1. Convergence of ADMM Iterates

$$\{(\mathbf{w}^k, \boldsymbol{\lambda}^k)\} \to (\mathbf{w}^*, \boldsymbol{\lambda}^*) \equiv \{\mathbf{v}^k\} \to \mathbf{v}^*$$

$$\begin{aligned} \mathbf{v}^k &= \mathbf{M}\mathbf{w}^k + (\mathbf{M} - \mathbf{I}_n)\boldsymbol{\lambda}^k - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b} \\ \mathbf{w}^{k+1} &= \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^k) \\ \boldsymbol{\lambda}^{k+1} &= \mathbb{P}_{\mathcal{Y}}(\mathbf{v}^k) - \mathbf{v}^k \end{aligned}$$

Let $(\mathbf{y}^*, \mathbf{w}^*, \boldsymbol{\lambda}^*)$ be optimal solution of (QP).
Denote, $\mathbf{v}^* = \mathbf{M}\mathbf{w}^* + (\mathbf{M} - \mathbf{I}_n)\boldsymbol{\lambda}^* - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b}$

ADMM - Convergence Analysis

2. Q-Linear Convergence of $\{\mathbf{v}^k\}$

$$\|\mathbf{v}^{k+1}-\mathbf{v}^*\| \leq \underbrace{\left(\sqrt{\|\tilde{\mathbf{M}}\|^2(1-\zeta^k)^2+\frac{1}{4}(\zeta^k)^2}+\frac{1}{2}\right)}_{\zeta^k<1,\ 0\prec\tilde{\mathbf{M}}\prec\mathbf{I}_{n-m}} \|\mathbf{v}^k-\mathbf{v}^*\|.$$

where, $\tilde{\mathbf{M}} = \frac{1}{2}(2\mathbf{Z}^T\mathbf{M}\mathbf{Z}-\mathbf{I}_{n-m})$,
$\zeta^k = \|\mathbf{R}\mathbf{R}^T\mathbf{u}^k\|/\|\mathbf{u}^k\|$, $\mathbf{u}^k = (2\mathbb{P}_{\mathcal{Y}}-\mathbf{I}_n)(\mathbf{v}^k)-(2\mathbb{P}_{\mathcal{Y}}-\mathbf{I}_n)(\mathbf{v}^*)$.

$$\alpha^0\|\mathbf{v}^0-\mathbf{v}^*\|, \alpha^0\alpha^1\|\mathbf{v}^0-\mathbf{v}^*\|, \cdots$$

Computational Certificate

Reach ϵ-optimal solution in: $O\left(\ln\left(\frac{1}{\epsilon}\right)\right)$ iterations

ADMM - Convergence Analysis

3. 2-Step Q-Linear Convergence of $\{(\mathbf{w}^k, \boldsymbol{\lambda}^k)\}$

$$\left\| \begin{pmatrix} \mathbf{w}^{k+2} \\ \boldsymbol{\lambda}^{k+2} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{*} \\ \boldsymbol{\lambda}^{*} \end{pmatrix} \right\| \le \alpha^k \left\| \begin{pmatrix} \mathbf{w}^{k} \\ \boldsymbol{\lambda}^{k} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{*} \\ \boldsymbol{\lambda}^{*} \end{pmatrix} \right\|$$

where, $\alpha^k = \left(\sqrt{\|\tilde{\mathbf{M}}\|^2 (1 - \zeta^k)^2 + \frac{1}{4}(\zeta^k)^2 + \frac{1}{2}} \right).$

Inequality 1: $\left\| \begin{pmatrix} \mathbf{w}^{k+2} \\ \boldsymbol{\lambda}^{k+2} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{*} \\ \boldsymbol{\lambda}^{*} \end{pmatrix} \right\| \le \|\mathbf{v}^{k+1} - \mathbf{v}^*\|$

Inequality 2: $\|\mathbf{v}^{k} - \mathbf{v}^*\| \le \left\| \begin{pmatrix} \mathbf{w}^{k} \\ \boldsymbol{\lambda}^{k} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{*} \\ \boldsymbol{\lambda}^{*} \end{pmatrix} \right\|$

Optimal Parameter Selection β^*

$$\|\mathbf{v}^{k+1} - \mathbf{v}^*\| \leq \left(\sqrt{\|\tilde{\mathbf{M}}\|^2 (1 - \zeta^k)^2 + \frac{1}{4} (\zeta^k)^2} + \frac{1}{2} \right) \|\mathbf{v}^k - \mathbf{v}^*\|$$

where, $\tilde{\mathbf{M}} = \frac{1}{2}(2\mathbf{Z}^T\mathbf{M}\mathbf{Z} - \mathbf{I}_{n-m})$,
No way to control ζ^k.

Choose β^* so as to maximize convergence rate

$$\beta^* = \arg\min_{\beta>0} \left(\frac{\|2\mathbf{Z}^T\mathbf{M}\mathbf{Z} - \mathbf{I}_{n-m}\|}{2} \right)$$

$$= \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\lambda_{\max}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})}$$

Intuition for $\beta^* = \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\lambda_{\max}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})}$

Solve for $\mathbf{y}^{k+1}$

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\xi^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

Intuition for $\beta^* = \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\lambda_{\max}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})}$

If $\beta << \beta^*$, then $\beta << \lambda_{\min}(\mathbf{Q})$.

$$\begin{aligned} \frac{1}{\beta}\mathbf{Q}\mathbf{y}^{k+1} + \frac{1}{\beta}\mathbf{A}^T\xi^{k+1} &\approx \boldsymbol{\lambda}^k - \frac{1}{\beta}\mathbf{q} \\ \mathbf{A}\mathbf{y}^{k+1} &= \mathbf{b} \end{aligned}$$

Emphasizing Optimality, or Decrease Dual Infeasibility

Solve for $\mathbf{y}^{k+1}$

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\xi^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

Intuition for $\beta^* = \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{QZ})\lambda_{\max}(\mathbf{Z}^T\mathbf{QZ})}$

If $\beta >> \beta^*$, then $\beta >> \lambda_{\max}(\mathbf{Q})$.

$$\begin{aligned} \mathbf{y}^{k+1} + \frac{1}{\beta}\mathbf{A}^T\xi^{k+1} &\approx \mathbf{w}^k \\ \mathbf{A}\mathbf{y}^{k+1} &= \mathbf{b} \end{aligned}$$

Emphasizing Feasibility, or Decrease Primal Inf

Solve for $\mathbf{y}^{k+1}$

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\xi^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

Intuition for $\beta^* = \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\lambda_{\max}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})}$

$$\beta^* = \sqrt{\lambda_{\min}(\mathbf{Q})\lambda_{\max}(\mathbf{Q})}$$

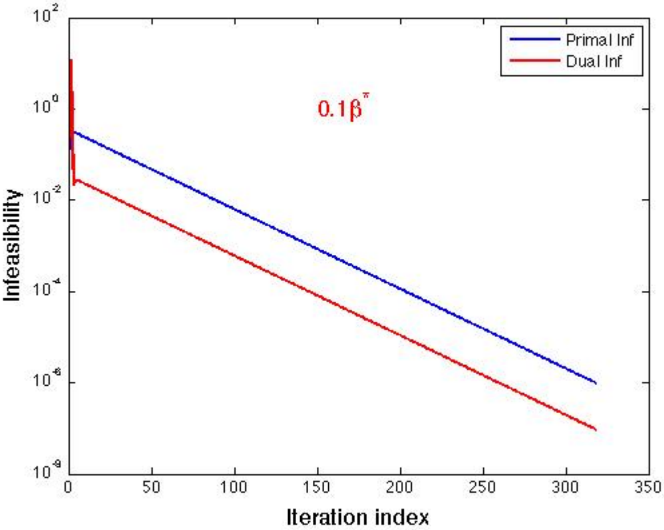
Balances both Primal and Dual Inf reduction automatically

Solve for $\mathbf{y}^{k+1}$

$$\begin{pmatrix} \mathbf{Q}/\beta + \mathbf{I}_n & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}^{k+1} \\ \frac{1}{\beta}\xi^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^k + \boldsymbol{\lambda}^k - \mathbf{q}/\beta \\ \mathbf{b} \end{pmatrix}$$

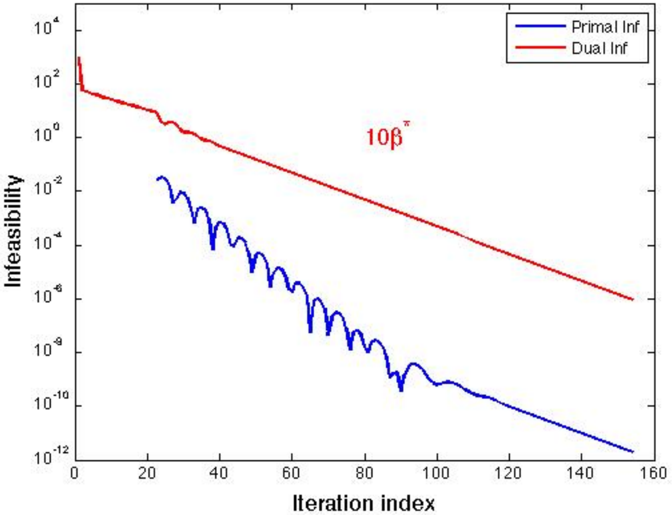
Results - Validating β^*

Small $\beta << \beta^*$



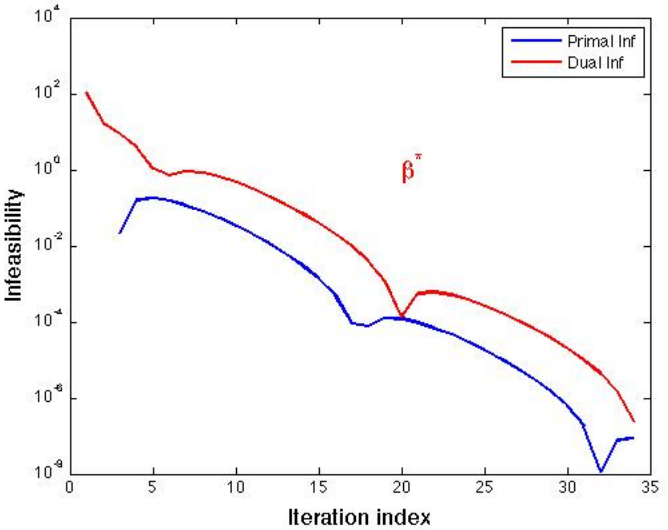
Results - Validating β^*

Large $\beta >> \beta^*$



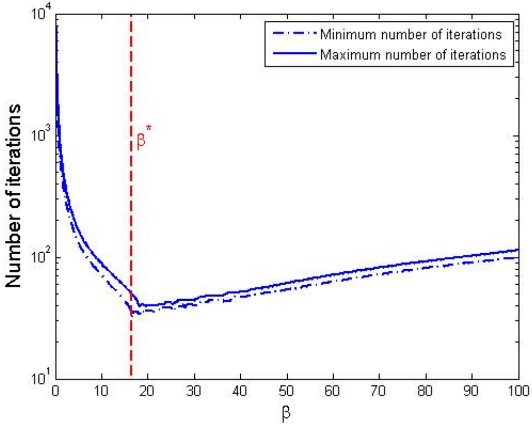
Results - Validating β^*

$$\beta = \beta^*$$



Results - Validating β^*

- 170 MPC QPs from Ghadimi et al
- All 170 QPs run at different value of β
- Max and min iterations across the 170 problems at each β



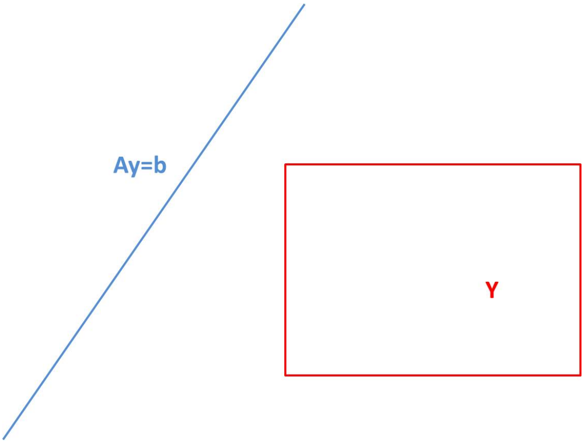
Outline

Infeasible (QP)

(QP) is infeasible

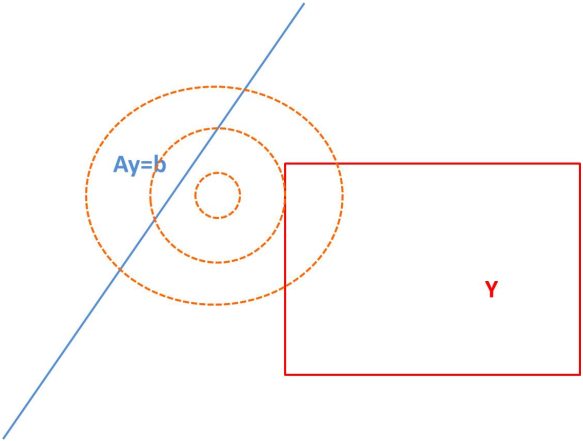
$$\Updownarrow$$

$$\{\mathbf{y} | \mathbf{A}\mathbf{y} = \mathbf{b}\} \cap \mathcal{Y} = \emptyset.$$



Infeasible (QP)

- What happens to ADMM iterates?
- No convergence of $\{\boldsymbol{\lambda}^k\}$: $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \mathbf{w}^{k+1} - \mathbf{y}^{k+1}$
- What is the influence of the objective function?

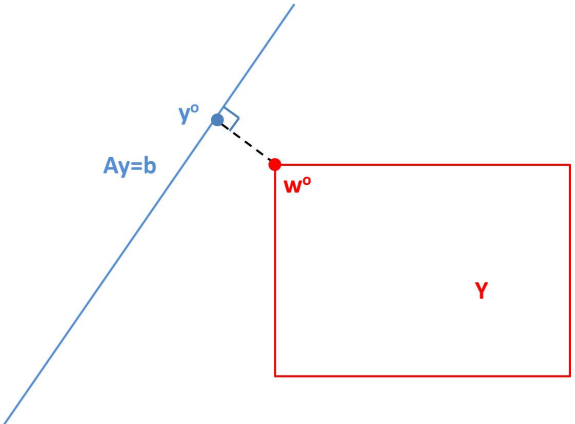


Infeasibility Minimizer

Let $(\mathbf{y}^\circ, \mathbf{w}^\circ)$ be optimal solution of

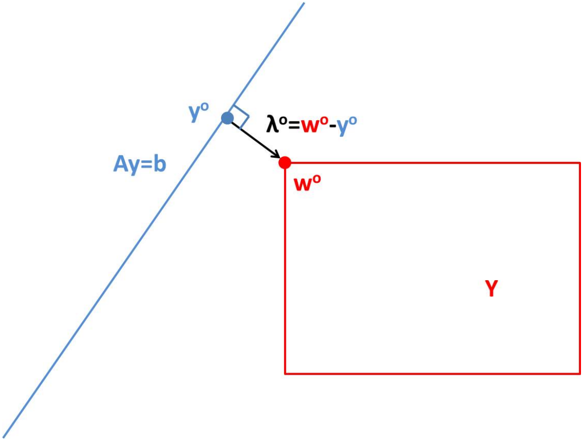
$$\min_{\mathbf{y},\mathbf{w}} \frac{1}{2} \|\mathbf{y} - \mathbf{w}\|^2$$

$$\text{s.t. } \mathbf{A}\mathbf{y} = \mathbf{b}, \mathbf{w} \in \mathcal{Y}$$



Infeasibility Minimizer

$$\begin{aligned} \mathbf{A}\mathbf{y}^\circ &= \mathbf{b} \\ \mathbf{w}^\circ - \mathbf{y}^\circ &= \boldsymbol{\lambda}^\circ \\ \mathbf{w}^\circ \in \mathcal{Y} &\perp \boldsymbol{\lambda}^\circ \end{aligned}$$

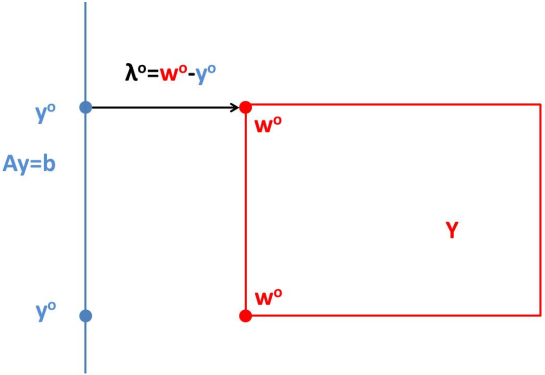


Non-unique Infeasibility Minimizer

- $\mathbf{y}^\circ, \mathbf{w}^\circ$ unique only in range$(\mathbf{A}^T)$

$$\mathbf{w}^\circ - \mathbf{y}^\circ \in \text{range}(\mathbf{A}^T)$$

$$\mathbf{w}^\circ \in \mathcal{Y} \perp \mathbf{w}^\circ - \mathbf{y}^\circ (= \boldsymbol{\lambda}^\circ)$$



Behavior of $\{\mathbf{v}^{k+1} - \mathbf{v}^k\}$

$$\mathbf{v}^{k+1} - \mathbf{v}^k = \mathbf{M}(\mathbf{w}^{k+1} - \mathbf{w}^k) + (\mathbf{M} - \mathbf{I}_n)(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k)$$

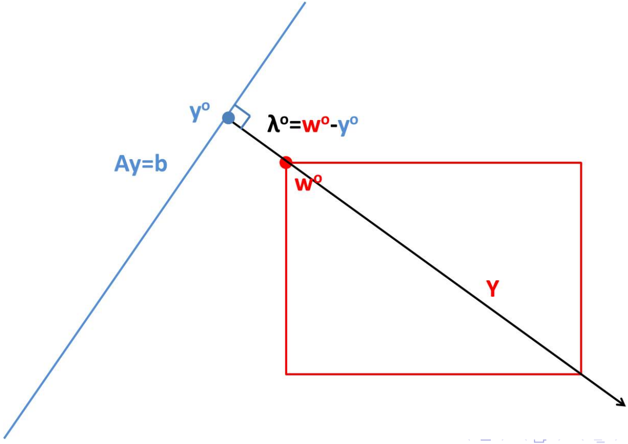
- $0 \preceq \mathbf{M} \prec \mathbf{I}_n \implies$ null space components of $\mathbf{w}^k, \boldsymbol{\lambda}^k$ are diminshed
- $\mathbf{M} = \mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z}/\beta + \mathbf{I}_{n-m})^{-1}\mathbf{Z}^T \implies$ range space components are irrelevant

$$\mathbf{v}^{k+1} - \mathbf{v}^k \to -(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k) \neq 0$$

ADMM Iterates Limiting Sequence (partial result)

(QP) is infeasible <u>if and only if</u> ADMM iterates satisfy:

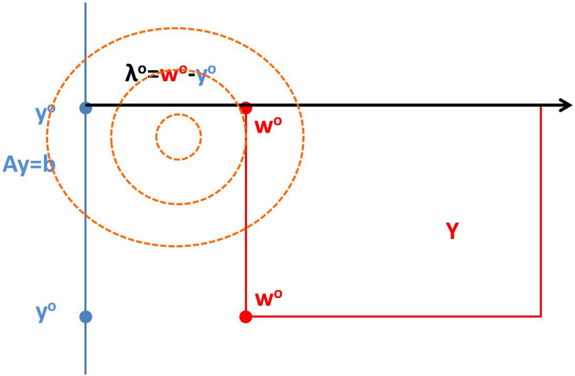
- $\{\mathbf{y}^k\} \to \mathbf{y}^\circ$
- $\{\mathbf{w}^k\} \to \mathbf{w}^\circ$
- $\{\boldsymbol{\lambda}^k\}$ satisfy $\boldsymbol{\lambda}^k = (k-\gamma)\boldsymbol{\lambda}^\circ$, $\gamma > 0$



ADMM Iterates Limiting Sequence

(QP) is infeasible if and only if ADMM iterates satisfy:

- $\{\mathbf{y}^k\} \to \mathbf{y}^\circ + \mathbf{y}^\mathbf{Q}$, where $\mathbf{y}^\mathbf{Q} \in \text{range}(\mathbf{Z})$
- $\{\mathbf{w}^k\} \to \mathbf{w}^\circ + \mathbf{y}^\mathbf{Q}$
- $\{\boldsymbol{\lambda}^k\}$ satisfy $\boldsymbol{\lambda}^k = \boldsymbol{\lambda}^\mathbf{Q} + (k-\gamma)\boldsymbol{\lambda}^\circ$, $\gamma > 0$



Null-Space Shift in ADMM Iterates

$\mathbf{y^Q} \in \text{range}(\mathbf{Z}), \boldsymbol{\lambda}^\mathbf{Q}$ satisfy

$$\mathbf{Z}^T\mathbf{Q}(\mathbf{y}^\circ + \mathbf{y^Q}) + \mathbf{Z}^T\mathbf{q} - \mathbf{Z}^T\boldsymbol{\lambda}^\mathbf{Q} = 0$$
$$(\mathbf{w}^\circ + \mathbf{y^Q}) \in \mathcal{Y} \perp \boldsymbol{\lambda}^\mathbf{Q}$$

Equivalently,

$\mathbf{y^Q} = \mathbf{Z}\mathbf{y}_\mathbf{Z}^\mathbf{Q}$ solves

$$\min_{\mathbf{y}_\mathbf{Z}^\mathbf{Q}} \frac{1}{2}(\mathbf{y}_\mathbf{Z}^\mathbf{Q})^T(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\mathbf{y}_\mathbf{Z}^\mathbf{Q} + (\mathbf{Z}^T\mathbf{q} + \mathbf{Z}^T\mathbf{Q}\mathbf{y}^\circ)^T\mathbf{y}_\mathbf{Z}^\mathbf{Q}$$
$$\text{s.t. } \mathbf{w}^\circ + \mathbf{Z}\mathbf{y}_\mathbf{Z}^\mathbf{Q} \in \mathcal{Y}.$$

Null-Space Shift in ADMM Iterates

$\mathbf{y^Q} = \mathbf{Z}\mathbf{y_Z^Q}$ solves

$$\min_{\mathbf{y_Z^Q}} \frac{1}{2}(\mathbf{y_Z^Q})^T(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\mathbf{y_Z^Q} + (\mathbf{Z}^T\mathbf{q} + \mathbf{Z}^T\mathbf{Q}\mathbf{y}^\circ)^T\mathbf{y_Z^Q}$$
$$\text{s.t. } \mathbf{w}^\circ + \mathbf{Z}\mathbf{y_Z^Q} \in \mathcal{Y}.$$

- Always feasible, choose $\mathbf{y_Z^Q} = 0 \implies$ Weak Slater's CQ
- Strictly convex program, $\mathbf{Z}^T\mathbf{Q}\mathbf{Z} \succ 0 \implies$ Unique $\mathbf{y_Z^Q}$
- Unique $\mathbf{Z}^T\boldsymbol{\lambda}^\mathbf{Q}$ follows from

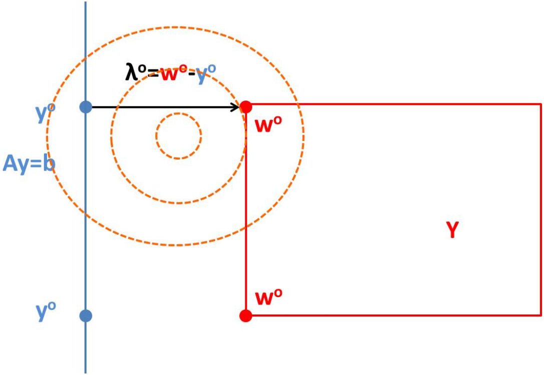
$$\mathbf{Z}^T\mathbf{Q}(\mathbf{y}^\circ + \mathbf{y^Q}) + \mathbf{Z}^T\mathbf{q} - \mathbf{Z}^T\boldsymbol{\lambda}^\mathbf{Q} = 0$$

- In fact,

$$\mathbf{w}^\circ + \mathbf{Z}\mathbf{y_Z^Q} \in \mathcal{Y} \perp \boldsymbol{\lambda}^\mathbf{Q} + \gamma_1\boldsymbol{\lambda}^\circ$$

Null-Space Shift in ADMM Iterates

- $\mathbf{y}_Z^{\mathbf{Q}} \neq 0$ such that $\mathbf{w}^\circ + \mathbf{Z}\mathbf{y}_Z^{\mathbf{Q}} \in \mathcal{Y}$ exist in these cases only
- Still choice of $\mathbf{Q}, \mathbf{q}$ influence if indeed $\mathbf{y}^{\mathbf{Q}} \neq 0$



ADMM - Convergence Analysis

1. Convergence of $\{\mathbf{v}^{k+1} - \mathbf{v}^k\}$ iterates

$$\|\mathbf{v}^{k+1} - \mathbf{v}^k\| = \|\mathbf{v}^k - \mathbf{v}^{k-1}\| \neq 0$$

holds iff $\mathbf{y}^k = \mathbf{y}^\circ + \mathbf{y}^{\mathbf{Q}}$, $\mathbf{w}^k = \mathbf{w}^\circ + \mathbf{y}^{\mathbf{Q}}$ and $\boldsymbol{\lambda}^k = \boldsymbol{\lambda}^{\mathbf{Q}} + (k - \gamma)\boldsymbol{\lambda}^\circ$.

ADMM - Convergence Analysis

2. Q-Linear Convergence of $\{\mathbf{v}^{k+1} - \mathbf{v}^k\}$

$$\|\mathbf{v}^{k+1} - \mathbf{v}^k\| \leq \underbrace{\left(\sqrt{\|\tilde{\mathbf{M}}\|^2(1-\zeta^k)^2 + \frac{1}{4}(\zeta^k)^2} + \frac{1}{2}\right)}_{\zeta^k<1,\ 0\prec\tilde{\mathbf{M}}\prec\mathbf{I}_{n-m}} \|\mathbf{v}^k - \mathbf{v}^{k-1}\|.$$

where, $\tilde{\mathbf{M}} = \frac{1}{2}(2\mathbf{Z}^T\mathbf{M}\mathbf{Z} - \mathbf{I}_{n-m})$,
$\zeta^k = \|\mathbf{R}\mathbf{R}^T\mathbf{u}^k\|/\|\mathbf{u}^k\|$, $\mathbf{u}^k = (2\mathbb{P}_{\mathcal{Y}} - \mathbf{I}_n)(\mathbf{v}^k) - (2\mathbb{P}_{\mathcal{Y}} - \mathbf{I}_n)(\mathbf{v}^{k-1})$.

$$\alpha^1\|\mathbf{v}^1 - \mathbf{v}^0\|, \alpha^1\alpha^2\|\mathbf{v}^1 - \mathbf{v}^0\|, \cdots$$

Computational Certificate

Reach ϵ–infeasible solution in: $O\left(\ln\left(\frac{1}{\epsilon}\right)\right)$ iterations

ADMM - Convergence Analysis

3. 2-Step Q-Linear Convergence of $\{(\mathbf{w}^k, \boldsymbol{\lambda}^k) - (\mathbf{w}^{k-1}, \boldsymbol{\lambda}^{k-1})\}$

$$\left\| \begin{pmatrix} \mathbf{w}^{k+2} \\ \boldsymbol{\lambda}^{k+2} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{k+1} \\ \boldsymbol{\lambda}^{k+1} \end{pmatrix} \right\| \leq \alpha^k \left\| \begin{pmatrix} \mathbf{w}^{k} \\ \boldsymbol{\lambda}^{k} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{k-1} \\ \boldsymbol{\lambda}^{k-1} \end{pmatrix} \right\|$$

where, $\alpha^k = \left(\sqrt{\|\tilde{\mathbf{M}}\|^2 (1 - \zeta^k)^2 + \frac{1}{4}(\zeta^k)^2} + \frac{1}{2} \right)$.

Inequality 1: $\left\| \begin{pmatrix} \mathbf{w}^{k+2} \\ \boldsymbol{\lambda}^{k+2} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{k+1} \\ \boldsymbol{\lambda}^{k+1} \end{pmatrix} \right\| \leq \|\mathbf{v}^{k+1} - \mathbf{v}^k\|$

Inequality 2: $\|\mathbf{v}^k - \mathbf{v}^{k-1}\| \leq \left\| \begin{pmatrix} \mathbf{w}^{k} \\ \boldsymbol{\lambda}^{k} \end{pmatrix} - \begin{pmatrix} \mathbf{w}^{k-1} \\ \boldsymbol{\lambda}^{k-1} \end{pmatrix} \right\|$

Optimal Parameter Selection β^*

$$\|\mathbf{v}^{k+1} - \mathbf{v}^k\| \leq \left(\sqrt{\|\tilde{\mathbf{M}}\|^2 (1-\zeta^k)^2 + \frac{1}{4}(\zeta^k)^2} + \frac{1}{2} \right) \|\mathbf{v}^k - \mathbf{v}^{k-1}\|$$

where, $\tilde{\mathbf{M}} = \frac{1}{2}(2\mathbf{Z}^T\mathbf{MZ} - \mathbf{I}_{n-m})$,
No way to control ζ^k.

Choose β^* so as to maximize convergence rate

$$\beta^* = \arg\min_{\beta > 0} \left(\frac{\|2\mathbf{Z}^T\mathbf{MZ} - \mathbf{I}_{n-m}\|}{2} \right)$$
$$= \sqrt{\lambda_{\min}(\mathbf{Z}^T\mathbf{QZ})\lambda_{\max}(\mathbf{Z}^T\mathbf{QZ})}$$

Termination Conditions

Define

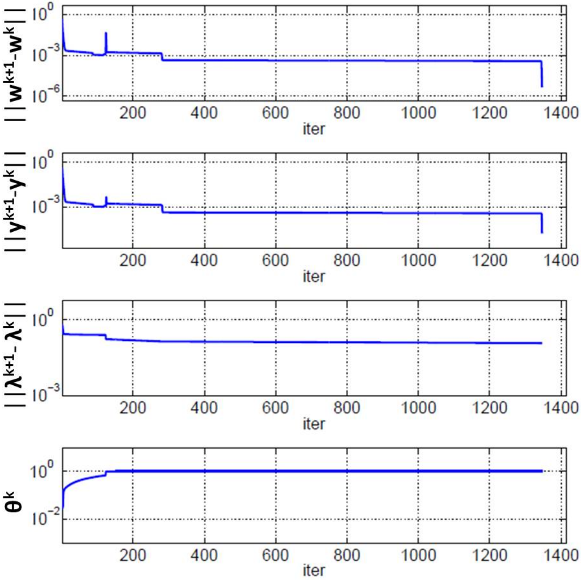
$$e_{opt}^k = \max(\beta\|\mathbf{w}^k - \mathbf{w}^{k-1}\|, \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}\|)$$
$$e_{inf}^k = \max(\beta\|\mathbf{w}^k - \mathbf{w}^{k-1}\|, \|\mathbf{y}^k - \mathbf{y}^{k-1}\|)$$
$$cos(\theta^k) = \frac{(\boldsymbol{\lambda}^k)^T(\mathbf{w}^k - \mathbf{y}^k)}{\|\boldsymbol{\lambda}^k\|\|\mathbf{w}^k - \mathbf{y}\|}$$

1. and 2. and 3. and (4. or 5.)

ϵ_i tolerance for infeasibility

1. $e_{opt} > \epsilon_o$ - tolerance for optimality
2. $e_{inf} \leq \epsilon_i e_{opt}$ - change in $\mathbf{w}, \mathbf{y} << $ error in optimality
3. $\theta^k \leq \epsilon_i$
4. $\|\mathbf{v}^k - \mathbf{v}^{k-1}\| \leq \epsilon_i \|\mathbf{v}^k\|$ - growth in magnitude of $\mathbf{v}^k$
5. $\boldsymbol{\lambda}_i^k(\mathbf{w}^k - \mathbf{y}^k)_i \geq 0$ - components of vectors are of same sign

Infeasible Instance



Outline

Solution of (QP)

- Suppose, $(\mathbf{y}^*, \boldsymbol{\lambda}^*)$ is optimal solution of (QP)
- Let $\mathcal{Y} = [\mathbf{y}^l, \mathbf{y}^u]$, $I^* := \{i | \mathbf{y}_i^* \text{ at the bounds}\}$
- $\mathbf{E}^*$ are gradients of $\mathbf{y}_i : i \in I^*$,
- $\bar{y}^*$ are the appropriate bound values

Given I^* solve

$$\begin{aligned} \min_{\mathbf{y}} \ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{E}^*\mathbf{y} = \bar{\mathbf{y}}^* \end{aligned} \tag{EQP*}$$

ADMM as Active-set Identifier

- Let $I^k(\epsilon) := \{i | \mathbf{y}_i^{k+1} \text{ within } \epsilon \text{ of bounds or } |\boldsymbol{\lambda}_i^{k+1}| > 0\}$
- Expect: $I^k(\epsilon) = I^*$ for all k sufficiently large

Given $I^k(\epsilon)$ solve

$$\begin{aligned} \min_{\mathbf{y}} \ & \mathbf{q}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q}\mathbf{y} \\ \text{s.t. } & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & \mathbf{E}^k\mathbf{y} = \bar{\mathbf{y}}^k \end{aligned} \tag{EQP}$$

where, rows of $\mathbf{E}^k$ are gradients of $\mathbf{y}_i : i \in I^k(\epsilon)$, $\bar{y}^k$ are the appropriate bound values.

Solving (EQP)

Substitute $\mathbf{y} = \mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{Z}\mathbf{y}_\mathbf{Z}$ with $\mathbf{y}_\mathbf{R} = (\mathbf{AR})^{-1}\mathbf{b}$ and simplifying,

Given $I^k(\epsilon)$ solve

$$\begin{aligned} \min_{\mathbf{y}_\mathbf{Z}} \; & (\mathbf{Z}^T\mathbf{Q}\mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{Z}^T\mathbf{q})^T\mathbf{y}_\mathbf{Z} + \frac{1}{2}\mathbf{y}_\mathbf{Z}^T(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\mathbf{y}_\mathbf{Z} \\ \text{s.t. } & \mathbf{E}^k\mathbf{Z}\mathbf{y}_\mathbf{Z} = \bar{\mathbf{y}}^k - \mathbf{E}^k\mathbf{R}\mathbf{y}_\mathbf{R} \end{aligned} \tag{rEQP}$$

Optimality conditions of (EQP) is:

$$\begin{bmatrix} \mathbf{Z}^T\mathbf{Q}\mathbf{Z} & -\mathbf{Z}^T(\mathbf{E}^k)^T \\ \mathbf{E}^k\mathbf{Z} & 0 \end{bmatrix} \begin{pmatrix} \mathbf{y}_\mathbf{Z} \\ \lambda_{I^k} \end{pmatrix} = \begin{pmatrix} -\mathbf{Z}^T(\mathbf{Q}\mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{q})(=: \mathbf{r}_1) \\ \bar{\mathbf{y}}^k - \mathbf{E}^k\mathbf{R}\mathbf{y}_\mathbf{R}(=: \mathbf{r}_2) \end{pmatrix}$$

Solving (EQP)

Substitute $\mathbf{y} = \mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{Z}\mathbf{y}_\mathbf{Z}$ with $\mathbf{y}_\mathbf{R} = (\mathbf{A}\mathbf{R})^{-1}\mathbf{b}$ and simplifying,

Given $I^k(\epsilon)$ solve

$$\begin{aligned} \min_{\mathbf{y}_\mathbf{Z}} \ & (\mathbf{Z}^T\mathbf{Q}\mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{Z}^T\mathbf{q})^T\mathbf{y}_\mathbf{Z} + \frac{1}{2}\mathbf{y}_\mathbf{Z}^T(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})\mathbf{y}_\mathbf{Z} \\ \text{s.t. } & \mathbf{E}^k\mathbf{Z}\mathbf{y}_\mathbf{Z} = \bar{\mathbf{y}}^k - \mathbf{E}^k\mathbf{R}\mathbf{y}_\mathbf{R} \end{aligned} \tag{rEQP}$$

Solution for $\boldsymbol{\lambda}_{I^k}$:

$$\left(\mathbf{E}^k\mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{E}^k)^T\right)\lambda_{I^k} = \mathbf{r}_2 - \mathbf{E}^k\mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})^{-1}\mathbf{r}_1$$

where, $\mathbf{r}_1 = -\mathbf{Z}^T(\mathbf{Q}\mathbf{R}\mathbf{y}_\mathbf{R} + \mathbf{q}), \mathbf{r}_2 = \bar{\mathbf{y}}^k - \mathbf{E}^k\mathbf{R}\mathbf{y}_\mathbf{R}$.

Solving (EQP) using CG

$$\left(\mathbf{E}^k\mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{E}^k)^T\right)\lambda_{I^k} = \mathbf{r}_2 - \mathbf{E}^k\mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})^{-1}\mathbf{r}_1$$

- $\mathbf{Z}^T\mathbf{Q}\mathbf{Z}$ is assumed to be positive definite
- If, $Z^T(\mathbf{E}^k)^T$ or equivalently, $[(\mathbf{E}^k)^T \;\; \mathbf{R}]$ has full column rank then,
  $$\mathbf{E}^k\mathbf{Z}(\mathbf{Z}^T\mathbf{Q}\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{E}^k)^T \succ 0$$
- CG can be applied to solve above system

Integrating ADMM w/ CG

- Difficult to gauge when $I^k(\epsilon) = I^*$. May only occur as $k \to \infty$
  - Call CG after every few ADMM iterations
  - Perform only few iterations of CG $< |I^k(\epsilon)|$

Integrating ADMM w/ CG

- Check if $\boldsymbol{\lambda}_{I^k}$ are of the correct sign
  - Indication that incorrect bounds were picked
  - If not, remove indices $i \in I^k(\epsilon)$ with incorrect multiplier sign
  - Perform CG on reduced set

Integrating ADMM w/ CG

- Has CG improved on ADMM
  - Use $\{\|\mathbf{v}^{k+1} - \mathbf{v}^k\|\}$ is non-increasing sequence
  - Perform 2 iterations of ADMM on CG solution
  - Verify that $\|\mathbf{v}^{\mathrm{cg},++} - \mathbf{v}^{\mathrm{cg},+}\| < \|\mathbf{v}^{k+1} - \mathbf{v}^k\|$
  - If yes, perform more CG iterations

ADMM with CG

Choose $\mathbf{w}^0, \boldsymbol{\lambda}^0, n^{\mathrm{cg,freq}}, n^{\mathrm{cg,max}}$. Set $k = 0$.

while termination condition not satisfied **do**

Set $\mathbf{v}^k = \mathbf{M}\mathbf{w}^k + (\mathbf{I} - \mathbf{M})\boldsymbol{\lambda}^k - \mathbf{M}\mathbf{q}/\beta + \mathbf{N}\mathbf{b}$

$(\mathbf{w}^{k+1}, \boldsymbol{\lambda}^{k+1}) = \mathrm{ADMM}(\mathbf{v}^k)$

Set $k = k + 1$

Every $n^{\mathrm{cg,freq}}$ ADMM iterations do

$(\mathbf{w}^{\mathrm{cg}}, \boldsymbol{\lambda}^{\mathrm{cg}}) = \mathrm{CG}(\mathbf{y}^k, \boldsymbol{\lambda}^k, I^k(\epsilon), n^{\mathrm{cg,max}})$

$(\mathbf{w}^{\mathrm{cg},+}, \boldsymbol{\lambda}^{\mathrm{cg},+}) = \mathrm{ADMM}(\mathbf{w}^{\mathrm{cg}}, \boldsymbol{\lambda}^{\mathrm{cg}})$

$(\mathbf{w}^{\mathrm{cg},++}, \boldsymbol{\lambda}^{\mathrm{cg},++}) = \mathrm{ADMM}(\mathbf{w}^{\mathrm{cg}}, \boldsymbol{\lambda}^{\mathrm{cg}})$

if $\|\mathbf{v}^{\mathrm{cg},++} - \mathbf{v}^{\mathrm{cg},+}\| \leq \|\mathbf{v}^k - \mathbf{v}^{k-1}\|$ **then**

Set $(\cdot)^k = (\cdot)^{\mathrm{cg},++}$, $(\cdot)^{k-1} = (\cdot)^{\mathrm{cg},++}$

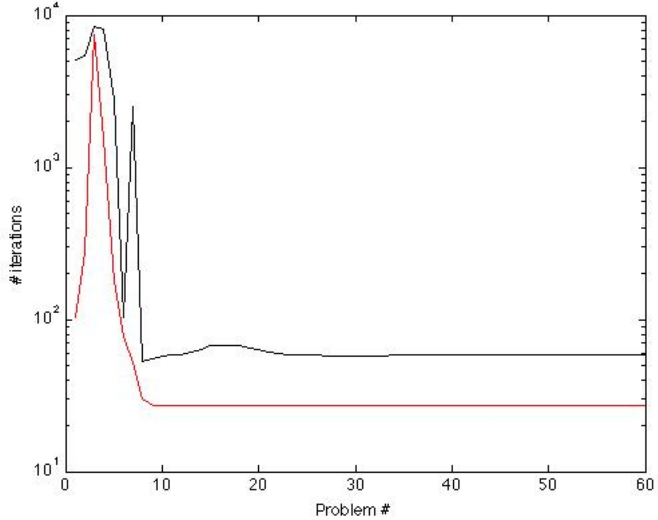
Go to CG call step

CUTEr Test Set ($n^{\mathrm{cg,max}} = 5$, $n^{\mathrm{cg,freq}} = 25$)

Problem	# vars	# eq cons	ADMM it	ADMM+CG it
DUAL1	85	1	89	42
DUAL2	96	1	127	78
CVXQP_1_S	100	50	34	29
CVXQP3_M	1000	750	35	29
CVXQP3_S	100	75	1548	66
AUG3D	3873	1000	45	33
AUG3DC	3873	1000	179	147
AUG3DCQP	3873	1000	1677	29
AUG3DQP	3873	1000	179	30
DTOC3	14999	9998	835	198
AUG2D	20200	10000	112	51
AUG2DC	20200	10000	80	43
AUG2DCQP	20200	10000	77	32
AUG2DQP	20200	10000	51	33

Results - MPC Test Problems

ADMM, ADMM+CG



Conclusions

- General convex QPs
  - Equality constrained QP + Simple projection operation
- Established 2-step Q-linear convergence rates for ADMM
- Derived optimal parameter β^*
- Handle Feasible and Infeasible cases
- Convergence acceleration
- Results generalize to general convex sets $\mathcal{Y}$
- Forthcoming: Decomposition methods for Stochastic MPC
  - Equality QP separable by scenarios
  - $\mathcal{Y}$ includes non-anticipativity- still **simple** projection

References

- **ACC 2014**: Alternating Direction Method of Multipliers (ADMM) for Strictly Convex Programs : Optimal Parameter Selection
- **MTNS 2014**: Optimal step-size selection in alternating direction method of multipliers for convex quadratic programs and model predictive control
- **CDC 2014**: Infeasibility Detection in Alternating Direction Method of Multipliers for Convex Quadratic Programs
- **Forthcoming**: Alternating Direction Method of Multipliers for General Convex Quadratic Programs: Optimal Parameter Selection, Infeasibility Detection and Converegence Acceleration
- **Forthcoming**: Decomposition Methods in ADMM for Scenario-based Stochastic MPC