Numerical Optimal Control with DAEs Lecture 6: Interior-Point Method

Sébastien Gros

AWESCO PhD course

1 / 21

Survival map of Direct Optimal Control



Survival map of Direct Optimal Control



Let's approach again the problem of solving the KKT conditions

2 / 21

< 🗗 > <

Outline



2 Primal Interior-Point Methods

3 Primal-Dual Interior-Point Methods

4 Primal-Dual Interior-Point Solver

5 Warm-start in Interior-Point Methods

KKT conditions - Reminder

Consider the NLP problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi\left(\mathbf{w}\right) \\ \text{s.t.} & \mathbf{g}\left(\mathbf{w}\right) = \mathbf{0} \\ & \mathbf{h}\left(\mathbf{w}\right) \leq \mathbf{0} \end{split}$$

KKT conditions with $\mathcal{L} = \Phi(\mathbf{w}) + \lambda^{\mathsf{T}} \mathbf{g}(\mathbf{w}) + \mu^{\mathsf{T}} \mathbf{h}(\mathbf{w})$

Primal Feasibility:	$\mathbf{g}\left(\mathbf{w} ight)=0,\mathbf{h}\left(\mathbf{w} ight)\leq0,$
Dual Feasibility:	$ abla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w}, oldsymbol{\mu}, oldsymbol{\lambda} ight)=0, oldsymbol{\mu}\geq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0, \forall i$

3

KKT conditions - Reminder

Consider the NLP problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi\left(\mathbf{w}\right) \\ \text{s.t.} & \mathbf{g}\left(\mathbf{w}\right) = \mathbf{0} \\ & \mathbf{h}\left(\mathbf{w}\right) \leq \mathbf{0} \end{split}$$

KKT conditions with $\mathcal{L} = \Phi(\mathbf{w}) + \lambda^{\mathsf{T}} \mathbf{g}(\mathbf{w}) + \mu^{\mathsf{T}} \mathbf{h}(\mathbf{w})$

Primal Feasibility:	$\mathbf{g}\left(\mathbf{w} ight)=0,\mathbf{h}\left(\mathbf{w} ight)\leq$	0,
Dual Feasibility:	$ abla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w}, oldsymbol{\mu}, oldsymbol{\lambda} ight)=0,$	$oldsymbol{\mu} \geq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0, \forall i$	

The difficulty of the KKT conditions is the non-smooth **Complementary Slackness** conditions resulting from the inequality constraints. Remember: "constraint \mathbf{h}_i can push $(\boldsymbol{\mu}_i > 0)$ only when \mathbf{w} touches it (i.e. when $\mathbf{h}_i = 0$)"

KKT conditions - Reminder

Consider the NLP problem:

 $\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$



Primal Feasibility:	$\mathbf{g}\left(\mathbf{w} ight)=0,\mathbf{h}\left(\mathbf{w} ight)\leq0,$
Dual Feasibility:	$ abla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w}, oldsymbol{\mu}, oldsymbol{\lambda} ight)=0, oldsymbol{\mu}\geq0,$
Complementary Slackness:	$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0, \forall i$

5

3

0

-5

3 2

The difficulty of the KKT conditions is the non-smooth **Complementary Slackness** conditions resulting from the inequality constraints. Remember: "constraint \mathbf{h}_i can push $(\boldsymbol{\mu}_i > 0)$ only when \mathbf{w} touches it (i.e. when $\mathbf{h}_i = 0$)"

 $\mathbf{h}_i(\mathbf{w})$ active

0

Solution manifold of $\mu_i \mathbf{h}_i(\mathbf{w}) = 0$

 $h_i(w)$ not active

 $h_i(w)$

-3 -2 -1



KKT conditions with $\mathcal{L} = \Phi(\mathbf{w}) + \lambda^{\mathsf{T}} \mathbf{g}(\mathbf{w}) + \mu^{\mathsf{T}} \mathbf{h}(\mathbf{w})$

Primal Feasibility:	$\mathbf{g}\left(\mathbf{w} ight)=0,\mathbf{h}\left(\mathbf{w} ight)\leq0,$
Dual Feasibility:	$ abla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},oldsymbol{\mu},oldsymbol{\lambda} ight)=0,oldsymbol{\mu}\geq0,$
Complementary Slackness:	$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0, \forall i$

The difficulty of the KKT conditions is the non-smooth **Complementary Slackness** conditions resulting from the inequality constraints. Remember: "constraint \mathbf{h}_i can push $(\mu_i > 0)$ only when \mathbf{w} touches it (i.e. when $\mathbf{h}_i = 0$)"



KKT conditions with $\mathcal{L} = \frac{1}{2}w^2 - w + \mu w$

Primal Feasibility:	$w \leq 0,$	
Dual Feasibility:	$w-1+\mu=0,$	$\mu\geq0,$
Complementary Slackness:	$\mu w = 0$	

The difficulty of the KKT conditions is the non-smooth **Complementary Slackness** conditions resulting from the inequality constraints. Remember: "constraint h_i can push $(\mu_i > 0)$ only when w touches it (i.e. when $h_i = 0$)"



KKT conditions with $\mathcal{L} = \frac{1}{2}w^2 - w + \mu w$

Primal Feasibility:	$w \leq 0,$	
Dual Feasibility:	$w-1+\mu=0,$	$\mu\geq 0,$
Complementary Slackness:	$\mu w = 0$	

The difficulty of the KKT conditions is the non-smooth **Complementary Slackness** conditions resulting from the inequality constraints. Remember: "constraint h_i can push $(\mu_i > 0)$ only when w touches it (i.e. when $h_i = 0$)"



KKT conditions with $\mathcal{L} = \frac{1}{2}w^2 - w + \mu w$

Primal Feasibility:	$w \leq 0,$	
Dual Feasibility:	$w-1+\mu=0,$	$\mu\geq 0,$
Complementary Slackness:	$\mu w = 0$	

Original idea of the IP method: introduce the inequality constraints in the cost !!

- 2

イロト イポト イヨト イヨト









Log-barrier method: introduce the inequality constraints in the cost function



Log-barrier approximates the characteristic function

$$\chi(\mathbf{h}_i) = \begin{cases} 0 & \text{if } \mathbf{h}_i \leq 0 \\ \infty & \text{if } \mathbf{h}_i > 0 \end{cases}$$



Log-barrier method: introduce the inequality constraints in the cost function

 $\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \min_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$

Example:

$$\min_{w} \quad \frac{1}{2}w^2 - 2w$$

s.t.
$$-1 \le w \le 1$$

.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \leq 0$$

6 / 21

過 ト イヨト イヨト

Log-barrier method: introduce the inequality constraints in the cost function

Example:

min w

$$\frac{1}{2}w^{2} - 2w$$

$$-1 \le w \le 1$$

$$\Phi_{\tau}(w) = \frac{1}{2}w^{2} - 2w - \tau \log(w+1) - \tau \log(1-w)$$

.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$h(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

Log-barrier method: introduce the inequality constraints in the cost function

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \\ \min\limits_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$$

Example:

Lee.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 1$$

Log-barrier method: introduce the inequality constraints in the cost function

$$\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{w} & becomes \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \min_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$$

Example:

$$\min_{w} \quad \frac{1}{2}w^{2} - 2w \text{s.t.} \quad -1 \le w \le 1$$

$$\Phi_{\tau}(w) = \frac{1}{2}w^{2} - 2w - \tau \log(w+1) - \tau \log(1-w) \tau = 0.51795$$

I.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 0.51795$$

Log-barrier method: introduce the inequality constraints in the cost function

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \min\limits_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$$

Example:

.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$h(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 0.26827$$

Log-barrier method: introduce the inequality constraints in the cost function

Example:

l.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 0.13895$$

Log-barrier method: introduce the inequality constraints in the cost function

$$\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \sup_{\mathbf{w}} & becomes \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \min_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$$

Example:

l.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \leq 0$$

$$\tau = 0.071969$$

Log-barrier method: introduce the inequality constraints in the cost function

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \sup\limits_{\mathbf{w},\tau} & becomes \end{array} & \min\limits_{\mathbf{w},\tau} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

Example:

l.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 0.037276$$

Log-barrier method: introduce the inequality constraints in the cost function

Example:

.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \leq 0$$

$$\tau = 0.019307$$

э

< 回 > < 三 > < 三

$$\min_{w} \quad \frac{1}{2}w^{2} - 2w \text{s.t.} \quad -1 \le w \le 1$$

$$\Phi_{\tau}(w) = \frac{1}{2}w^{2} - 2w - \tau \log(w+1) - \tau \log(1-w) \tau = 0.01$$

I.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$h(w) = \begin{bmatrix} -w - 1 \\ w - 1 \end{bmatrix} \le 0$$

$$\tau = 0.01$$

Log-barrier method: introduce the inequality constraints in the cost function

Φ

 $\begin{array}{ll} \min\limits_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \\ \min\limits_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$

Example:

$$\min_{w} \quad \frac{1}{2}w^2 - 2w$$

s.t.
$$-1 \le w \le 1$$

.e.
$$\Phi(w) = \frac{1}{2}w^2 - 2w$$
$$\mathbf{h}(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \leq 0$$

$$\tau_{\tau}(w) = \frac{1}{2}w^2 - 2w - \tau \log(w+1) - \tau \log(1-w)$$



Log-barrier method: introduce the inequality constraints in the cost function

Φ

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \mathrm{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array} \qquad \qquad \min\limits_{\mathbf{w}_{\tau}} \Phi_{\tau}\left(\mathbf{w}_{\tau}\right) = \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}_{\tau})) \end{array}$$

Example:

$$\min_{w} \quad \frac{1}{2}w^2 - 2w$$

s.t.
$$-1 \le w \le 1$$

.e.

$$\Phi(w) = \frac{1}{2}w^2 - 2w$$

$$h(w) = \begin{bmatrix} -w - 1\\ w - 1 \end{bmatrix} \le 0$$

If \mathbf{w}^* is LICQ & SOSC, then $\|\mathbf{w}^*_{ au} - \mathbf{w}^*\| = \mathcal{O}(au)$

$$_{\tau}(w) = \frac{1}{2}w^2 - 2w - \tau \log(w+1) - \tau \log(1-w)$$



Problem:

 $\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{array}$

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}))$$



- 4 周 ト - 4 日 ト - 4 日

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ s.t. & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) &= 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad \mu \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}))$$



▲ □ ▶ ▲ □ ▶ ▲ □

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$\begin{aligned} \nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} &= 0 \\ \boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) &= 0 \\ \mathbf{h}(\mathbf{w}) &\leq 0, \quad \boldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}))$$

m

KKT conditions*:

$$abla \Phi_{ au}\left(\mathbf{w}
ight) =
abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_{t}} \mathbf{h}_{i}(\mathbf{w})^{-1}
abla \mathbf{h}_{i}(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$



A (1) > A (2) > A

Problem:

$$\label{eq:phi} \begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & \mu_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad \mu \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}))$$

KKT conditions*:

$$abla \Phi_{ au}\left(\mathbf{w}
ight) =
abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_{t}} \mathbf{h}_{i}(\mathbf{w})^{-1}
abla \mathbf{h}_{i}(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Newton direction for the Primal Interior-Point KKTs:



・ 同 ト ・ ヨ ト ・ ヨ

Problem:

$$\label{eq:phi} \begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}))$$

KKT conditions*:

$$abla \Phi_{ au}\left(\mathbf{w}
ight) =
abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_{t}} \mathbf{h}_{i}(\mathbf{w})^{-1}
abla \mathbf{h}_{i}(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Newton direction for the Primal Interior-Point KKTs:

$$\left(\nabla^{2}\Phi(\mathbf{w}) + \tau \sum_{i=1}^{m_{i}} \mathbf{h}_{i}(\mathbf{w})^{-2} \nabla \mathbf{h}_{i} \nabla \mathbf{h}_{i}^{\mathsf{T}}\right) \Delta \mathbf{w} + \Phi_{\tau}\left(\mathbf{w}\right) = 0$$

for ${\bf h}$ affine



Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & \mu_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad \mu \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}} \Phi_{\tau}(\mathbf{w}) = \min_{\mathbf{w}} \Phi(\mathbf{w}) - \tau \sum_{i=1}^{m_{i}} \log(-\mathbf{h}_{i}(\mathbf{w}))$$

KKT conditions*:

$$abla \Phi_{ au}\left(\mathbf{w}
ight) =
abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_{t}} \mathbf{h}_{i}(\mathbf{w})^{-1}
abla \mathbf{h}_{i}(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Newton direction for the Primal Interior-Point KKTs:

$$\left(\nabla^{2}\Phi(\mathbf{w}) + \tau \sum_{i=1}^{m_{i}} \mathbf{h}_{i}(\mathbf{w})^{-2} \nabla \mathbf{h}_{i} \nabla \mathbf{h}_{i}^{\mathsf{T}}\right) \Delta \mathbf{w} + \Phi_{\tau}\left(\mathbf{w}\right) = 0$$

for ${\bf h}$ affine

As $\tau \to 0$, the term $\mathbf{h}_i^{-2}(\mathbf{w})$ becomes very large when $\mathbf{h}_i \to 0$, which hinders the convergence



Primal-Dual Interior-Point method

Problem: $\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$ KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \mu = 0$ $\mu_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \mu \geq 0$

過す イヨト イヨト

- 34

Primal-Dual Interior-Point method

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) &= 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w})
abla \mathbf{h}_i(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

(日) (圖) (E) (E) (E)
Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) &= 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w})
abla \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

(日) (同) (三) (三) (三)

- 2

Problem:

 $\Phi(\mathbf{w})$ min 337 s.t. $h(w) \leq 0$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Introduce variable $\nu_i = -\tau \mathbf{h}_i^{-1}(\mathbf{w})$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w})
abla \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Problem:

min $\Phi(\mathbf{w})$ s.t. h(w) < 0

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w})
abla \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Introduce variable $\nu_i = -\tau \mathbf{h}_i^{-1}(\mathbf{w})$, then the Primal-Dual KKT conditions[†] read as: $abla \Phi(\mathbf{w}) + \sum_{i=1}^{m_i} \boldsymbol{\nu}_i \nabla \mathbf{h}_i(\mathbf{w}) = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) = -\tau$ [†]valid for $\mathbf{h}_i(\mathbf{w}) < 0, \ \boldsymbol{\nu}_i > 0$

8 / 21

Problem:

 $\Phi(\mathbf{w})$ min s.t. h(w) < 0

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w})
abla \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Introduce $\nu_i = -\tau \mathbf{h}_i^{-1}$, then the Primal-Dual Interior-Point KKT conditions read as: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0}$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0}$ $h(w) < 0, \nu > 0$

Problem:

 $\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ s.t. & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$

KKT conditions:

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

Barrier formulation:

$$\min_{\mathbf{w}_{\tau}} \Phi(\mathbf{w}_{\tau}) - \tau \sum_{i=1}^{m_i} \log(-\mathbf{h}_i(\mathbf{w}_{\tau}))$$

KKT conditions*:

$$abla \Phi(\mathbf{w}) - au \sum_{i=1}^{m_i} \mathbf{h}_i^{-1}(\mathbf{w}) \nabla \mathbf{h}_i(\mathbf{w}) = 0$$

*valid for $\mathbf{h}_i(\mathbf{w}) < 0$

Introduce $\nu_i = -\tau \mathbf{h}_i^{-1}$, then the Primal-Dual Interior-Point KKT conditions read as: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \nu = 0$ $\nu_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \nu > 0$

Primal-Dual IP conditions yield the same solution as the Barrier problem

Observe the similitude with the original KKT conditions !!

S. Gros

Optimal Control with DAEs, lecture 6

17th of February, 2016

3

8 / 21

(日) (四) (日) (日) (日)



 $\mathbf{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \mathbf{\nu} > 0$

- 3

A D A D A A D A

9 / 21

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = \mathbf{0}$ $\mu_i \mathbf{h}_i(\mathbf{w}) = 0$ $h(w) \leq 0, \quad \mu > 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0}$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0}$ $h(w) < 0, \nu > 0$



- 3

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \boldsymbol{\tau} = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



A

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = \mathbf{0}$ $\mu_i \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$ $h(w) \leq 0, \quad \mu > 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0}$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0}$ $h(w) < 0, \nu > 0$



- 3

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \boldsymbol{\tau} = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \boldsymbol{\tau} = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



A 1

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = \mathbf{0}$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$ $\mathbf{h}(\mathbf{w}) \le \mathbf{0}, \quad \boldsymbol{\mu} \ge \mathbf{0}$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \boldsymbol{\tau} = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = \mathbf{0}$ $\mu_i \mathbf{h}_i(\mathbf{w}) = \mathbf{0}$ $h(w) \leq 0, \quad \mu > 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0}$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0}$ $h(w) < 0, \nu > 0$



- 3

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

Primal-Dual IP KKT conditions $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\nu} = 0$ $\boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$ $\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\nu} > 0$



• Primal-Dual IP method solves KKT conditions with smoothed complementary slackness

(4 間) トイヨト イヨト

Problem:

$$\begin{array}{ll} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

 $\begin{aligned} & \mathsf{Primal-Dual \ IP \ KKT \ conditions} \\ & \nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0} \\ & \boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0} \\ & \mathbf{h}(\mathbf{w}) < \mathbf{0}, \quad \boldsymbol{\nu} > \mathbf{0} \end{aligned}$



- Primal-Dual IP method solves KKT conditions with smoothed complementary slackness
- IP approximation

$$\|\boldsymbol{\mu}^* - \boldsymbol{\nu}^*\| = \mathcal{O}(\tau)$$
$$\|\mathbf{w}^* - \mathbf{w}^*_{\tau}\| = \mathcal{O}(\tau)$$

(4 間) トイヨト イヨト

 $\mathbf{w}^*_{ au}$, $oldsymbol{
u}^*$ and \mathbf{w}^* , $oldsymbol{\mu}^*$ are equivocated

Problem:

$$\begin{split} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{h}(\mathbf{w}) \leq \mathbf{0} \end{split}$$

KKT conditions: $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\mu} = 0$ $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) \leq 0, \quad \boldsymbol{\mu} \geq 0$

 $\begin{aligned} & \mathsf{Primal-Dual \ IP \ KKT \ conditions} \\ & \nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w}) \boldsymbol{\nu} = \mathbf{0} \\ & \boldsymbol{\nu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0} \\ & \mathbf{h}(\mathbf{w}) < \mathbf{0}, \quad \boldsymbol{\nu} > \mathbf{0} \end{aligned}$

Note: the PD-IP KKT conditions require that **w** is **inside** the feasible domain



- Primal-Dual IP method solves KKT conditions with smoothed complementary slackness
- IP approximation

$$\|\boldsymbol{\mu}^* - \boldsymbol{\nu}^*\| = \mathcal{O}(\tau)$$
$$\|\mathbf{w}^* - \mathbf{w}^*_{\tau}\| = \mathcal{O}(\tau)$$

 $\mathbf{w}^*_{ au}$, $oldsymbol{
u}^*$ and \mathbf{w}^* , $oldsymbol{\mu}^*$ are equivocated

NLP $\begin{array}{c} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$

KKT conditions

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{g}(\mathbf{w}) oldsymbol{\lambda} +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0 \ & \mathbf{g}(\mathbf{w}) = 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) &= 0 \ & \mathbf{h}(\mathbf{w}) \leq 0, \quad oldsymbol{\mu} \geq 0 \end{aligned}$$

 17^{th} of February, 2016 10 / 21

- 31

(4月) (4日) (4日)

NLP

$$\begin{array}{c} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

PD-IP KKT conditions

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{g}(\mathbf{w}) oldsymbol{\lambda} +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathbf{0} \ & \mathbf{g}(\mathbf{w}) = \mathbf{0} \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = \mathbf{0} \ & \mathbf{h}(\mathbf{w}) < \mathbf{0}, \quad oldsymbol{\mu} > \mathbf{0} \end{aligned}$$

 $17^{
m th}$ of February, 2016

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

3

10 / 21

NLP

$$\begin{array}{c} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

PD-IP KKT conditions

 $egin{aligned}
abla \mathcal{L}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) &= 0 \ & \mathbf{g}\left(\mathbf{w}
ight) &= 0 \ & oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au &= 0 \ & \mathbf{h}(\mathbf{w}) < 0, \quad oldsymbol{\mu} > 0 \end{aligned}$

- 3

・ 母 ト ・ ヨ ト ・ ヨ ト

$$\begin{aligned} \textbf{JLP} \\ \min_{\mathbf{w}} & \Phi\left(\mathbf{w}\right) \\ \text{s.t.} & \mathbf{g}\left(\mathbf{w}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w}\right) \leq 0 \end{aligned}$$

Newton on the conditions (parametrized by τ)

$$\begin{bmatrix} \nabla \mathcal{L} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g} (\mathbf{w}) \\ \boldsymbol{\mu}_i \mathbf{h}_i (\mathbf{w}) + \tau \end{bmatrix} = \mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0}$$

with $\mathbf{h}(\mathbf{w}) < 0$, $\mu > 0$

17th of February, 2016 10 / 21

- 2

イロト イポト イヨト イヨト

NLP

$$\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0 \\ \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

Newton on the conditions (parametrized by τ)

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g}(\mathbf{w}) \\ \boldsymbol{\mu}_{i} \mathbf{h}_{i}(\mathbf{w}) + \tau \end{bmatrix} = \mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0}$$

with $\mathbf{h}(\mathbf{w}) < 0$, $\mu > 0$

Newton direction d given by

$$abla \mathbf{r}_{ au}^{ op} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \left[egin{array}{c} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \boldsymbol{\mu} \end{array}
ight] + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) = \mathbf{0}$$

Newton step: updates

$$\left[\begin{array}{c} \mathbf{w}\\ \boldsymbol{\lambda}\\ \boldsymbol{\mu} \end{array}\right] \leftarrow \left[\begin{array}{c} \mathbf{w}\\ \boldsymbol{\lambda}\\ \boldsymbol{\mu} \end{array}\right] + t \left[\begin{array}{c} \Delta \mathbf{w}\\ \Delta \boldsymbol{\lambda}\\ \Delta \boldsymbol{\mu} \end{array}\right]$$

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ● □ ● ○ ○ ○

NLP

$$\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0 \\ \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

Newton on the conditions (parametrized by τ)

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g}(\mathbf{w}) \\ \boldsymbol{\mu}_{i} \mathbf{h}_{i}(\mathbf{w}) + \tau \end{bmatrix} = \mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0}$$

with $\mathbf{h}(\mathbf{w}) < 0$, $\mu > 0$

Newton direction d given by

$$\nabla \mathbf{r}_{\tau}^{\top} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \boldsymbol{\lambda} \\ \Delta \boldsymbol{\mu} \end{bmatrix} + \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) = \mathbf{0}$$

Newton step: updates

$$\left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] \leftarrow \left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] + t \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \end{array}
ight]$$

Step-size: $t \in]0, 1]$ must ensure:

 $h(w + t\Delta w) < 0, \quad \mu + t\Delta \mu > 0$

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ● □ ● ○ ○ ○

NLP $\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0 \\ \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$ **Newton on the conditions** (parametrized by τ)

$$\begin{bmatrix} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) \\ \mathbf{g} \left(\mathbf{w} \right) \\ \boldsymbol{\mu}_{i} \mathbf{h}_{i} (\mathbf{w}) + \tau \end{bmatrix} = \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) = \mathbf{0}$$

with $\mathbf{h}(\mathbf{w}) < 0$, $\mu > 0$

Newton direction d given by

$$abla \mathbf{r}_{ au}^{ op}\left(\mathbf{w},oldsymbol{\lambda},oldsymbol{\mu}
ight) \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \end{array}
ight] + \mathbf{r}_{ au}\left(\mathbf{w},oldsymbol{\lambda},oldsymbol{\mu}
ight) = \mathbf{0}$$

Newton step: updates

$$\left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] \leftarrow \left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] + t \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \end{array}
ight]$$

Step-size: $t \in]0, 1]$ must ensure:

 $h(w + t\Delta w) < 0, \quad \mu + t\Delta \mu > 0$

Difficulties:

• Selecting t to get

 $h(w + t\Delta w) < 0$

cannot be done simply. Requires evaluating **h** for decreasingly large values of *t* until the condition is met. Can be expensive !!

17th of February, 2016

NLP

$$\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0 \\ \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$$

Newton on the conditions (parametrized by τ)

$$\begin{bmatrix} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) \\ \mathbf{g} \left(\mathbf{w} \right) \\ \boldsymbol{\mu}_{i} \mathbf{h}_{i} (\mathbf{w}) + \tau \end{bmatrix} = \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu} \right) = \mathbf{0}$$

with $\mathbf{h}(\mathbf{w}) < 0$, $\mu > 0$

Newton direction d given by

$$abla \mathbf{r}_{ au}^{ op}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta oldsymbol{\mu} \end{array}
ight] + \mathbf{r}_{ au}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) = \mathbf{0}$$

Newton step: updates

$$\left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] \leftarrow \left[egin{array}{c} \mathbf{w} \ \lambda \ \mu \end{array}
ight] + t \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \end{array}
ight]$$

Step-size: $t \in]0, 1]$ must ensure:

 $h(w + t\Delta w) < 0, \quad \mu + t\Delta \mu > 0$

Difficulties:

• Selecting t to get

 $h(w + t\Delta w) < 0$

cannot be done simply. Requires evaluating \mathbf{h} for decreasingly large values of tuntil the condition is met. Can be expensive !!

• We need the initial guess to be feasible for h !!

10 / 21

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathbf{0}$$
 (1a)

$$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + \tau = \mathbf{0}$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

- Newton steps on (1a)-(1b)
- Backtrack to ensure (1c)

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0$$
 (1a)

$$\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = \mathbf{0}$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

- Newton steps on (1a)-(1b)
- Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

11 / 21

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathbf{0}$$
 (1a)

$$oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = \mathbf{0}$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

- Newton steps on (1a)-(1b)
- Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

need a feasible initial guess

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathbf{0}$$
 (1a)

$$oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = \mathbf{0}$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

- Newton steps on (1a)-(1b)
- Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

- need a feasible initial guess
- backtracking can be expensive

Primal-Dual IP KKT conditions:		Slack reformulation: new variable -
$ abla \Phi(\mathbf{w}) + abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0$ (1a))	$ abla \Phi(\mathbf{w}) + abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0$
$oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = 0$ (1b))	$\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$
$\mathbf{h}(\mathbf{w}) < 0, \boldsymbol{\mu} > 0$ (10)	:)	$-oldsymbol{\mu}_i oldsymbol{s}_i + au = 0$
		$-\mathbf{s} < 0, oldsymbol{\mu} > 0$
Newton steps on (1a)-(1b)		
Backtrack to ensure (1c)		
 Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations need a feasible initial guess backtracking can be expensive 		

э

(人間) トイヨト イヨト

 $-\mathbf{s} = \mathbf{h}(\mathbf{w})$

Primal-Dual IP KKT conditions:		Slack reformulation: new variable $-\mathbf{s} = \mathbf{h}\left(\mathbf{w}\right)$
$ abla \Phi(\mathbf{w}) + abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0$	(1a)	$ abla \Phi(\mathbf{w}) + abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0$
$oldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}) + au = 0$	(1b)	$\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$
$\mathbf{h}(\mathbf{w}) < 0, oldsymbol{\mu} > 0$	(1c)	$oldsymbol{\mu}_i oldsymbol{s}_i - au = oldsymbol{0}$
		$\mathbf{s} > 0, \boldsymbol{\mu} > 0$
Newton steps on (1a)-(1b)	
 Backtrack to ensure (1c) 		
 Difficulty: one must ensure that h(w) starts and remains negati throughout the Newton iteratio need a feasible initial gues backtracking can be expension 	nt ve ns ss nsive	

э

- 4 同 6 4 日 6 4 日 6

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathbf{0}$$
 (1a)

$$\mu_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

- need a feasible initial guess
- backtracking can be expensive

Slack reformulation: new variable $-\mathbf{s} = \mathbf{h}(\mathbf{w})$ $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = 0$ $\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$ $\boldsymbol{\mu}_i s_i - \tau = 0$ $\mathbf{s} > 0, \quad \boldsymbol{\mu} > 0$

Newton on the slack formulation

• initialize with ${f s},\, {m \mu}>0$ and ${m \mu}_i {m s}_i= au$

超す イヨト イヨト

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathsf{0}$$
 (1a)

$$\mu_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

- need a feasible initial guess
- backtracking can be expensive

Slack reformulation: new variable $-\mathbf{s} = \mathbf{h}(\mathbf{w})$ $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = 0$ $\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$ $\boldsymbol{\mu}_i s_i - \tau = 0$ $\mathbf{s} > 0, \quad \boldsymbol{\mu} > 0$

Newton on the slack formulation

- initialize with ${f s},\, {m \mu}>0$ and ${m \mu}_i {m s}_i= au$
- h(w) > 0 does not matter at the initial guess or during the iterations

< 回 > < 三 > < 三 >

Primal-Dual IP KKT conditions:

$$abla \Phi(\mathbf{w}) +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = \mathsf{0}$$
 (1a)

$$\mu_i \mathbf{h}_i(\mathbf{w}) + \tau = 0$$
 (1b)

$$\mathbf{h}(\mathbf{w}) < 0, \quad \boldsymbol{\mu} > 0$$
 (1c)

Backtrack to ensure (1c)

Difficulty: one must ensure that h(w) starts and remains negative throughout the Newton iterations

- need a feasible initial guess
- backtracking can be expensive

Slack reformulation: new variable $-\mathbf{s} = \mathbf{h}(\mathbf{w})$ $\nabla \Phi(\mathbf{w}) + \nabla \mathbf{h}(\mathbf{w})\boldsymbol{\mu} = 0$ $\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$ $\boldsymbol{\mu}_i s_i - \tau = 0$ $\mathbf{s} > 0, \quad \boldsymbol{\mu} > 0$

Newton on the slack formulation

- initialize with ${f s},\, {m \mu}>0$ and ${m \mu}_i {m s}_i= au$
- h(w) > 0 does not matter at the initial guess or during the iterations
- finding $t \in [0, 1]$ to enforce:

 $\mathbf{s} + t\Delta \mathbf{s} > 0$ $\boldsymbol{\mu} + t\Delta \boldsymbol{\mu} > 0$

is trivial.

KKT conditions with slack

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{g}(\mathbf{w}) oldsymbol{\lambda} +
abla \mathbf{h}(\mathbf{w}) oldsymbol{\mu} = 0 \ \mathbf{g}(\mathbf{w}) = 0 \ \mathbf{h}(\mathbf{w}) + \mathbf{s} = 0 \ \mu_i oldsymbol{s}_i = 0 \ \mathbf{s} > 0, \quad oldsymbol{\mu} > 0 \end{aligned}$$

3

PD-IP KKT conditions with slack

$$egin{aligned}
abla \Phi(\mathbf{w}) +
abla \mathbf{g}(\mathbf{w}) \lambda +
abla \mathbf{h}(\mathbf{w}) \mu &= 0 \ \mathbf{g}(\mathbf{w}) &= 0 \ \mathbf{h}(\mathbf{w}) + \mathbf{s} &= 0 \ \mu_i s_i - au &= 0 \ \mathbf{s} &\geq 0, \quad \mu &\geq 0 \end{aligned}$$

 $17^{
m th}$ of February, 2016
PD-IP KKT conditions

$$egin{aligned}
abla \mathcal{L}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) &= 0 \ \mathbf{g}\left(\mathbf{w}
ight) &= 0 \ \mathbf{h}(\mathbf{w}) + \mathbf{s} &= 0 \ oldsymbol{\mu}_i oldsymbol{s}_i - au &= 0 \ \mathbf{s} &> 0, \quad oldsymbol{\mu} &> 0 \end{aligned}$$

э

$$\begin{aligned} \text{ILP} \\ \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & h(\mathbf{w}) \leq 0 \end{aligned}$$

Ν

Newton on the conditions

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g}(\mathbf{w}) \\ \mathbf{h}(\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_i s_i - \tau \end{bmatrix} = \mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

with s > 0, $\mu > 0$

 $17^{
m th}$ of February, 2016

- 2

12 / 21

(日) (同) (三) (三) (三)

NLP $\begin{array}{l} \min_{\mathbf{w}} \quad \Phi(\mathbf{w}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0 \\ \quad \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$

Newton on the conditions

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g}(\mathbf{w}) \\ \mathbf{h}(\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_{i} s_{i} - \tau \end{bmatrix} = \mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$
with $\mathbf{s} > 0, \quad \boldsymbol{\mu} > \mathbf{0}$

Newton direction d given by $abla \mathbf{r}_{\tau}^{\top}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) \mathbf{d} + \mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) = \mathbf{0}$

NLP $\min_{\mathbf{w}} \Phi(\mathbf{w})$ w s.t. $\mathbf{g}(\mathbf{w}) = \mathbf{0}$ $\mathbf{h}(\mathbf{w}) < 0$

Newton on the conditions

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \mathbf{g}(\mathbf{w}) \\ \mathbf{h}(\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_{i} \mathbf{s}_{i} - \tau \end{bmatrix} = \mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$
with $\mathbf{s} > 0$, $\boldsymbol{\mu} > 0$

Newton direction d given by $\nabla \mathbf{r}_{\tau}^{\top}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) \mathbf{d} + \mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$

$$\begin{array}{c|cccc} & \mathcal{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} & \operatorname{diag}(\mathbf{s}) & \operatorname{diag}(\boldsymbol{\mu}) \end{array} \end{bmatrix} \underbrace{ \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix}}_{=\mathbf{v}_{\mathsf{T}_{\tau}}(\mathbf{w}, \lambda, \mu, \mathbf{s})} = -\mathbf{r}_{\tau} \left(\mathbf{w}, \lambda, \mu, \mathbf{s} \right)$$
with $\mathcal{H} = \nabla^{2} \mathcal{L} \left(\mathbf{w}, \lambda, \mu \right)$

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ● □ ● ○ ○ ○ 17th of February, 2016

12 / 21

NLP $\begin{array}{c} \min_{\mathbf{w}} & \Phi(\mathbf{w}) \\ \text{s.t.} & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{array}$

Newton on the conditions

$$\begin{bmatrix} \nabla \mathcal{L}(\mathbf{w}, \lambda, \mu) \\ \mathbf{g}(\mathbf{w}) \\ \mathbf{h}(\mathbf{w}) + \mathbf{s} \\ \mu_{i} \mathbf{s}_{i} - \tau \end{bmatrix} = \mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = \mathbf{0}$$
with $\mathbf{s} > 0$, $\mu > 0$

17th of February, 2016

12 / 21

Newton direction d given by $\nabla \mathbf{r}_{\tau}^{\top} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s} \right) \mathbf{d} + \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s} \right) = \mathbf{0}$

$$\underbrace{\begin{bmatrix} H & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} & \operatorname{diag}(\mathbf{s}) & \operatorname{diag}(\boldsymbol{\mu}) \end{bmatrix}}_{=\nabla \mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s})} \underbrace{\begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix}}_{=\mathbf{d}} = -\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s})$$

with $H = \nabla^2 \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu})$

Observe the specific structure of the matrix $\nabla \mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s})$!!

Optimal Control with DAEs, lecture

Solve:

$$\mathbf{r}_{ au}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) \ \mathbf{g}\left(\mathbf{w}
ight) \ \mathbf{h}\left(\mathbf{w}
ight) + \mathbf{s} \ oldsymbol{\mu}_{i} s_{i} - au \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$



3

過す イヨト イヨト



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$ for a very small τ .

▲□▶▲圖▶▲≣▶▲≣▶ = ● ● ●

13 / 21

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g} \left(\mathbf{w}
ight) \\ \mathbf{h} (\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_{i} s_{i} - \tau \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$ for a very small τ .

Solve:

$$\mathbf{r}_{ au}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) \\
\mathbf{g}\left(\mathbf{w}
ight) \\
\mathbf{h}(\mathbf{w}) + \mathbf{s} \\
oldsymbol{\mu}_{i} \mathbf{s}_{i} - au \end{array}
ight] = 0$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g} \left(\mathbf{w}
ight) \\ \mathbf{h} \left(\mathbf{w}
ight) + \mathbf{s} \\ \boldsymbol{\mu}_{i} \mathbf{s}_{i} - \boldsymbol{\tau} \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Key idea: solve at large τ , then reduce it while solving again...

13 / 21

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\
\mathbf{g} \left(\mathbf{w}
ight) \\
\mathbf{h} \left(\mathbf{w}
ight) + \mathbf{s} \\
eta_{\boldsymbol{\mu}_{i}} \mathbf{s}_{i} - \mathbf{\tau} \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) = \left[egin{array}{c} \nabla \mathcal{L}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g}\left(\mathbf{w}
ight) \\ \mathbf{h}(\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_{i} s_{i} - au \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{ au}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\
\mathbf{g}\left(\mathbf{w}
ight) \\
\mathbf{h}(\mathbf{w}) + \mathbf{s} \\
eta_{i} s_{i} - au \end{array}
ight] = 0$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{ au}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L}\left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}
ight) \\
\mathbf{g}\left(\mathbf{w}
ight) \\
\mathbf{h}(\mathbf{w}) + \mathbf{s} \\
oldsymbol{\mu}_{i} s_{i} - au \end{array}
ight] = 0$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g} \left(\mathbf{w}
ight) \\ \mathbf{h} \left(\mathbf{w}
ight) + \mathbf{s} \\ \boldsymbol{\mu}_{i} s_{i} - \tau \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g} \left(\mathbf{w}
ight) \\ \mathbf{h} (\mathbf{w}) + \mathbf{s} \\ \boldsymbol{\mu}_{i} s_{i} - \tau \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c}
abla \mathcal{L}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\
\mathbf{g}\left(\mathbf{w}
ight) \\
\mathbf{h}(\mathbf{w}) + \mathbf{s} \\
eta_{i} s_{i} - au \end{array}
ight] = 0$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, oldsymbol{\lambda}, oldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Solve:

$$\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \left[egin{array}{c} \nabla \mathcal{L} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}
ight) \\ \mathbf{g} \left(\mathbf{w}
ight) \\ \mathbf{h} \left(\mathbf{w}
ight) + \mathbf{s} \\ \boldsymbol{\mu}_{i} \mathbf{s}_{i} - \boldsymbol{\tau} \end{array}
ight] = \mathbf{0}$$

Taking steps along the ...

Newton direction: d given by

$$abla \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight)^{ op} \mathbf{d} + \mathbf{r}_{ au} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}
ight) = \mathbf{0}$$

Reminder: Newton convergence depends on the Lipschitz constant of $\nabla \mathbf{r}_{\tau}$ ($\mathbf{w}, \lambda, \mu, \mathbf{s}$), i.e. Newton does not "like" strong nonlinearities



We want to solve $\mathbf{r}_{\tau}(\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$ for a very small τ . But we do not want to get through the "corner" in $\mu_i \mathbf{s}_i = \tau$ when τ is small.

Key idea: solve at large τ , then reduce it while solving again...

13 / 21

Key idea:

 Algorithm: PD-IP solver

 Set τ , μ , $\mathbf{s} \leftarrow 1$, guess \mathbf{w}, λ

 while $\tau > \operatorname{tol} \mathbf{do}$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$ with $0 < \gamma < 1$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$

Key idea:

 Algorithm: PD-IP solver

 Set τ , μ , $\mathbf{s} \leftarrow 1$, guess \mathbf{w}, λ

 while $\tau > \operatorname{tol} \mathbf{do}$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$ with $0 < \gamma < 1$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$



Key idea:

Example min $\frac{1}{2}(\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}}Q(\mathbf{w} - \mathbf{w}_0)$

s.t.
$$\mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1$$



14 / 21

Key idea:

Algorithm: PD-IP solver Set τ , μ , $\mathbf{s} \leftarrow 1$, guess $\mathbf{w}, \boldsymbol{\lambda}$ while $\tau > tol do$ Solve $\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s} \right) = 0$ Update $\tau \leftarrow \gamma \tau$ with $0 < \gamma < 1$ return w, λ , μ , s

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$



Key idea:

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$



Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, s) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, s$

 $\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$

Central path: solution manifold of

$$\mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) = \mathbf{0}$$



Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, s) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, s$

$$\begin{array}{l} \begin{array}{l} \begin{array}{l} \text{xample} \\ \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array} \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

for $\tau \in [1, 0[$

E





Key idea: homotopy on τ

Algorithm: PD-IP solver Set τ , μ , $\mathbf{s} \leftarrow 1$ while $\tau > tol do$ Solve $\mathbf{r}_{\tau} (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$ Update $\tau \leftarrow \gamma \tau$ return w, λ , μ , s

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

for $\tau \in [1, 0[$



Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$

for $\tau \in [1, 0[$



Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

for $\tau \in [1, 0[$



Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$



$$\gamma = 0.25$$

Key idea: homotopy on τ

 Algorithm: PD-IP solver

 Set $\tau, \mu, s \leftarrow 1$

 while $\tau > tol do$

 Solve $\mathbf{r}_{\tau} (\mathbf{w}, \lambda, \mu, \mathbf{s}) = 0$

 Update $\tau \leftarrow \gamma \tau$

 return $\mathbf{w}, \lambda, \mu, \mathbf{s}$

$$\begin{array}{ll} \textbf{Example} & & \\ & \min_{\mathbf{w}} & \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^{\mathsf{T}} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ & \text{s.t.} & \mathbf{w}^{\mathsf{T}} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

for $\tau \in [1, 0[$



Key idea: path-following

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$egin{aligned} & \prod\limits_{\mathbf{w}}^{\mathsf{T}} & rac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0
ight) \ & ext{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{aligned}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$

for $\tau \in [1, 0[$



Key idea: path-following

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$





 $\gamma = 0.1$

Key idea: path-following

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) = \mathbf{0}$$



Key idea: path-following

 $\begin{array}{l} \textbf{Algorithm: PD-IP solver} \\ \textbf{Set } \tau, \ \boldsymbol{\mu}, \ \mathbf{s} \leftarrow 1 \\ \textbf{while } \tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do} \\ \textbf{Newton step on } \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) \\ \textbf{if } \|\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right)\|_{X} \leq 1 \text{ then} \\ \textbf{L} \quad \textbf{Update } \tau \leftarrow \gamma \tau \end{array}$

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$





Key idea: path-following

 $\begin{array}{l} \textbf{Algorithm: PD-IP solver} \\ \textbf{Set } \tau, \ \boldsymbol{\mu}, \ \mathbf{s} \leftarrow 1 \\ \textbf{while } \tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do} \\ \textbf{Newton step on } \mathbf{r}_{\tau} \ (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) \\ \textbf{if } \|\mathbf{r}_{\tau} \ (\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s})\|_{X} \leq 1 \ \textbf{then} \\ \textbf{L} \ \textbf{Update } \tau \leftarrow \gamma \tau \end{array}$

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$





Key idea: path-following

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$




Solving an NLP using the Primal-Dual Interior-Point method

Key idea: path-following

 $\begin{array}{l} \textbf{Algorithm: PD-IP solver} \\ \textbf{Set } \tau, \ \boldsymbol{\mu}, \ \mathbf{s} \leftarrow 1 \\ \textbf{while } \tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do} \\ \textbf{Newton step on } \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) \\ \textbf{if } \|\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right)\|_{X} \leq 1 \text{ then} \\ \textbf{L} \quad \textbf{Update } \tau \leftarrow \gamma \tau \end{array}$

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \mathbf{0}$$

for $\tau \in [1, 0[$





Solving an NLP using the Primal-Dual Interior-Point method

Key idea: path-following

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = 0$$

for $\tau \in [1, 0[$





Solving an NLP using the Primal-Dual Interior-Point method

Key idea: path-following

 $\begin{array}{l} \textbf{Algorithm: PD-IP solver} \\ \textbf{Set } \tau, \ \boldsymbol{\mu}, \ \mathbf{s} \leftarrow 1 \\ \textbf{while } \tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do} \\ \textbf{Newton step on } \mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) \\ \textbf{if } \|\mathbf{r}_{\tau} \left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right)\|_{X} \leq 1 \text{ then} \\ \textbf{L} \quad \textbf{Update } \tau \leftarrow \gamma \tau \end{array}$

return $\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}$

Example

$$\begin{array}{l} \displaystyle \min_{\mathbf{w}} \quad \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^\mathsf{T} Q \left(\mathbf{w} - \mathbf{w}_0 \right) \\ \mathrm{s.t.} \quad \mathbf{w}^\mathsf{T} S \mathbf{w} \leq 1 \end{array}$$

Central path: solution manifold of

$$\mathbf{r}_{\tau}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\right) = \mathbf{0}$$

for $\tau \in [1, 0[$





Algorithm: a Primal-dual Interior-Point solver

Input: w Set $\tau = 1$, $\mu = 1$, $\mathbf{s} = 1$, $\lambda = 0$ while $\tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do}$ w, λ , μ , s return



S. Gros

 17^{th} of February, 2016

15 / 21

Algorithm: a Primal-dual Interior-Point solver





17th of February, 2016



17th of February, 2016

15 / 21



return w, λ , μ , s





return w, λ , μ , s



return w, λ , μ , s

Algorithm: a Primal-dual Interior-Point solver Input: w Set $\tau = 1$, $\mu = 1$, s = 1, $\lambda = 0$ $\mu_i \mathbf{s}_i = 0$ and $\mu_i \mathbf{s}_i = \tau$ while $\tau > \text{tol or } \|\mathbf{r}_{\tau}\|_{\infty} > \text{tol do}$ Evaluate H, g, h, ∇g , ∇h , $\nabla \Phi$ Compute the Newton direction given by $\dot{\mathbf{n}}^2$ $\begin{array}{c|cccc} H & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & I \\ \end{array} \begin{vmatrix} \Delta \mathbf{w} \\ \Delta \mathbf{\lambda} \\ \Delta \mu \\ \Delta \mathbf{r} \end{vmatrix} = -\mathbf{r}_{\tau}$ diag(s) diag(μ) Compute a step-size $t_{\text{max}} \leq 1$ ensuring: Some subtleties: $\mathbf{s} + t_{\max} \Delta \mathbf{s} \geq \epsilon \mathbf{s}, \quad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} \geq \epsilon \boldsymbol{\mu}$ Measuring progress Backtrack $t \in [0, t_{max}]$ to ensure progress • Choice of $\|.\|_X$ Take Newton step: $\mathbf{w} \leftarrow \mathbf{w} + t\Delta \mathbf{w}, \dots$ Mehrotra predictor if $\|\mathbf{r}_{\tau}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s})\|_{\mathrm{X}} < 1$ then Update $\tau \leftarrow \gamma \tau$ • "Adaptive" γ w, λ , μ , s return

A B F A B F

 μ_i

15 / 21

$$\begin{split} \min_{\mathbf{x},\,\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \, \|\mathbf{x}_{k} - x_{\mathrm{ref}}\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \, \|\mathbf{u}_{k} - u_{\mathrm{ref}}\|_{R}^{2} \\ \mathrm{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k}\right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

э

$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - x_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \le \mathbf{u} \le \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \le v \le v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2:$ forward acceleration

S. Gros

$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\text{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\text{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\text{max}} \le \mathbf{u} \le \mathbf{u}_{\text{max}}, \qquad -v_{\text{min}} \le v \le v_{\text{max}} \end{split}$$

Simple plane dynamics

Γ ×́ Τ		$v \cos(\theta)$
ý		$v \sin(\theta)$
$\dot{ heta}$	$= \mathbf{F} =$	$\mathit{gv}^{-1} \operatorname{tan}(\phi)$
$\dot{\phi}$		\mathbf{u}_1
, v		\mathbf{u}_2

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \le \mathbf{u} \le \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \le v \le v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\text{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\text{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\text{max}} \le \mathbf{u} \le \mathbf{u}_{\text{max}}, \qquad -v_{\text{min}} \le v \le v_{\text{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\text{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\text{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\text{max}} \le \mathbf{u} \le \mathbf{u}_{\text{max}}, \qquad -v_{\text{min}} \le v \le v_{\text{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\text{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\text{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \ge r^{2} \\ & - \mathbf{u}_{\text{max}} \le \mathbf{u} \le \mathbf{u}_{\text{max}}, \qquad -v_{\text{min}} \le v \le v_{\text{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi :$ bank angle
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros



$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \sum_{k=0}^{N} \frac{1}{2} \left\| \mathbf{x}_{k} - \mathbf{x}_{\mathrm{ref}} \right\|_{Q}^{2} + \sum_{k=0}^{N-1} \frac{1}{2} \left\| \mathbf{u}_{k} - u_{\mathrm{ref}} \right\|_{R}^{2} \\ \text{s.t} \quad & \mathbf{x}_{k+1} = \mathbf{x}_{k} + \Delta t \mathbf{F} \left(\mathbf{x}_{k}, \mathbf{u}_{k} \right) \\ & \mathbf{x}_{0} = \hat{\mathbf{x}}, \qquad x^{2} + y^{2} \geq r^{2} \\ & - \mathbf{u}_{\mathrm{max}} \leq \mathbf{u} \leq \mathbf{u}_{\mathrm{max}}, \qquad -v_{\mathrm{min}} \leq v \leq v_{\mathrm{max}} \end{split}$$

Simple plane dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \\ \dot{v} \end{bmatrix} = \mathbf{F} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ g v^{-1} \tan(\phi) \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

- x, y: position
- v: forward velocity
- θ : heading
- $\phi : \text{ bank angle}$
- $\mathbf{u}_1\text{: roll rate}$
- $\mathbf{u}_2\text{: forward acceleration}$

S. Gros





NMPC



NMPC

 $17^{
m th}$ of February, 2016 17 / 21



NMPC

 $17^{
m th}$ of February, 2016 $\,$ 17 / 21



NMPC



NMPC

 $17^{
m th}$ of February, 2016 17 / 21



S. Gros

Optimal Control with DAEs, lecture 6

17th of February, 2016 17 / 21



 $17^{
m th}$ of February, 2016 $\,$ 17 / 21



S. Gros

Optimal Control with DAEs, lecture 6

17th of February, 2016 1

17 / 21



NMPC


NMPC



Optimal Control with DAEs, lecture 6

NMPC



 $17^{
m th}$ of February, 2016 17 / 21





S. Gros

 $17^{
m th}$ of February, 2016 $\,$ 17 / 21



17th of February, 2016 17 / 21



NMPC

 17^{th} of February, 2016 17 / 21



NMPC

 17^{th} of February, 2016 17 / 21



NMPC

 17^{th} of February, 2016 17 / 21

Sparsity of the Primal-Dual Interior-Point KKT matrix:

$$\begin{bmatrix} H & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} & \operatorname{diag}(\mathbf{s}) & \operatorname{diag}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$

18 / 21

Sparsity of the Primal-Dual Interior-Point KKT matrix:

$$\begin{bmatrix} \boldsymbol{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \mathbf{0} & \mathsf{diag}(\mathbf{s}) & \mathsf{diag}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$

$$egin{bmatrix} H &
abla \mathbf{g} &
abla \mathbf{h} & 0 \
abla \mathbf{g}^{ op} & 0 & 0 & 0 \
abla \mathbf{h}^{ op} & 0 & 0 & I \ 0 & 0 & ext{diag}\left(\mathbf{s}
ight) & ext{diag}\left(\mathbf{\mu}
ight) \end{bmatrix} egin{bmatrix} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{ au}$$



S. Gros

 $17^{
m th}$ of February, 2016

$$\left[egin{array}{cccc} m{H} &
abla \mathbf{g} &
abla \mathbf{h} & 0 \
abla \mathbf{g}^{\mathsf{T}} & 0 & 0 & 0 \
abla \mathbf{h}^{\mathsf{T}} & 0 & 0 & I \ 0 & 0 & \mathrm{diag}\left(\mathbf{s}
ight) & \mathrm{diag}\left(\mathbf{\mu}
ight) \end{array}
ight] \left[egin{array}{c} \Delta \mathbf{w} \ \Delta \lambda \ \Delta \mu \ \Delta \mathbf{s} \end{array}
ight] = -\mathbf{r}_{ au}$$



Required ordering:

$$\begin{bmatrix} \boldsymbol{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \mathbf{0} & \mathsf{diag}\left(\mathbf{s}\right) & \mathsf{diag}\left(\boldsymbol{\mu}\right) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \boldsymbol{\mu} \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$



Required ordering:

$$\mathbf{g}\left(\mathbf{w}\right) = \left[\begin{array}{c} \mathbf{x}_{0} - \hat{\mathbf{x}} \\ \mathbf{f}\left(\mathbf{x}_{0}, \mathbf{u}_{0}\right) - \mathbf{x}_{1} \\ \\ \dots \\ \mathbf{f}\left(\mathbf{x}_{N-1}, \mathbf{u}_{N-1}\right) - \mathbf{x}_{N} \end{array} \right],$$

17th of February, 2016 18 / 21

$$\begin{bmatrix} \boldsymbol{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \mathbf{0} & \mathsf{diag}(\mathbf{s}) & \mathsf{diag}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$



18 / 21

$$\begin{bmatrix} \boldsymbol{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \mathbf{0} & \mathsf{diag}(\mathbf{s}) & \mathsf{diag}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$



$$\begin{bmatrix} \boldsymbol{H} & \nabla \mathbf{g} & \nabla \mathbf{h} & \mathbf{0} \\ \nabla \mathbf{g}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}^{\mathsf{T}} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \mathbf{0} & \mathsf{diag}(\mathbf{s}) & \mathsf{diag}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \lambda \\ \Delta \mu \\ \Delta \mathbf{s} \end{bmatrix} = -\mathbf{r}_{\tau}$$





... what happens if we have a very good guess to warm-start our algorithm ?



 $17^{
m th}$ of February, 2016







... what happens if we have a very good guess to warm-start our algorithm ?



 17^{th} of February, 2016





... what happens if we have a very good guess to warm-start our algorithm ?



 17^{th} of February, 2016

... what happens if we have a very good guess to warm-start our algorithm ?





Even with an excellent initial guess interior point methods will **retreat to the central path** before homing onto the solution... what about keeping τ low ?







... what happens if we have a very good guess to warm-start our algorithm ?



19 / 21
















... what happens if we have a very good guess to warm-start our algorithm ?



17th of February, 2016

19 / 21





At very low τ , changes of active set are difficult: Newton struggles to get through the sharp turn in $\mu_i s_i = \tau$



Survival map of Direct Optimal Control



Survival map of Direct Optimal Control



17th of February, 2016

Algorithm: SQP (prototype)

end

イロト イポト イヨト イヨト

- 2

Algorithm: SQP (prototype)

while Not converged do
Form
$$\nabla_{\mathbf{w}}^{2}\mathcal{L}$$
, $\nabla_{\mathbf{w}}\mathcal{L}$, \mathbf{g} , $\nabla \mathbf{g}$, \mathbf{h} , $\nabla \mathbf{h}$
while IPQP not converged do
Newton step on:
 $H\Delta \mathbf{w} + \nabla \Phi + \nabla \mathbf{g} \lambda^{\mathrm{QP}} + \nabla \mathbf{h} \mu^{\mathrm{QP}} = 0$
 $\nabla \mathbf{g}^{\mathrm{T}} \Delta \mathbf{w} + \mathbf{g} = 0$
 $\nabla \mathbf{h}^{\mathrm{T}} \Delta \mathbf{w} + \mathbf{h} + \mathbf{s}^{\mathrm{QP}} = 0$
 $\mu_{i}^{\mathrm{QP}} \mathbf{s}_{i}^{\mathrm{QP}} = \tau$
reduce $\tau \to \epsilon$
end
Update
 $\{\mathbf{w}, \lambda, \mu\} \leftarrow \{\mathbf{w}, \lambda, \mu\} + \Delta \{\mathbf{w}, \lambda, \mu\}$
end

イロト イポト イヨト イヨト

- 2

Algorithm: SQP (prototype)	Algorithm: IP (prototype)
while Not converged do Form $\nabla_{\mathbf{w}}^2 \mathcal{L}$, $\nabla_{\mathbf{w}} \mathcal{L}$, \mathbf{g} , $\nabla \mathbf{g}$, \mathbf{h} , $\nabla \mathbf{h}$ while IPQP not converged do Newton step on: $H\Delta \mathbf{w} + \nabla \Phi + \nabla \mathbf{g} \lambda^{\text{QP}} + \nabla \mathbf{h} \mu^{\text{QP}} = 0$ $\nabla \mathbf{g}^{\text{T}} \Delta \mathbf{w} + \mathbf{g} = 0$ $\nabla \mathbf{h}^{\text{T}} \Delta \mathbf{w} + \mathbf{h} + \mathbf{s}^{\text{QP}} = 0$ $\mu_i^{\text{QP}} \mathbf{s}_i^{\text{QP}} = \tau$	while Not converged do Form $\nabla^2_{\mathbf{w}} \mathcal{L}$, $\nabla_{\mathbf{w}} \mathcal{L}$, \mathbf{g} , $\nabla_{\mathbf{g}}$, \mathbf{h} , $\nabla \mathbf{h}$ Newton step on: $\nabla \mathcal{L}(\mathbf{w}, \lambda, \mu) = 0$ $\mathbf{g}(\mathbf{w}) = 0$ $\mathbf{h}(\mathbf{w}) + \mathbf{s} = 0$ $\mu_i s_i = \tau$ Update
$\left \begin{array}{c} reduce \ \tau \to \epsilon \\ end \\ Update \\ \{\mathbf{w}, \boldsymbol{\lambda}, \mu\} \leftarrow \{\mathbf{w}, \boldsymbol{\lambda}, \mu\} + \Delta \left\{\mathbf{w}, \boldsymbol{\lambda}, \mu\right\} \\ end \end{array} \right.$	$\begin{aligned} \{\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}\} \leftarrow \{\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}\} + \\ & \Delta \left\{\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right\} \\ \text{reduce } \tau \to \epsilon \\ \text{end} \end{aligned}$

- 2

・ロト ・聞ト ・ヨト ・ヨト

Algorithm: SQP (prototype)	Algorithm: IP (prototype)
$ \begin{array}{c c} \mbox{while Not converged do} \\ \mbox{Form } \nabla^2_{\mathbf{w}}\mathcal{L}, \ \nabla_{\mathbf{w}}\mathcal{L}, \ \mathbf{g}, \ \nabla \mathbf{g}, \ \mathbf{h}, \ \nabla \mathbf{h} \\ \mbox{while } IPQP \ not \ converged \ \mathbf{do} \\ \mbox{Newton step on:} \\ \mbox{H}\Delta \mathbf{w} + \nabla \Phi + \nabla \mathbf{g} \lambda^{\mathrm{QP}} + \nabla \mathbf{h} \mu \\ \mbox{\nabla} \mathbf{g}^{\mathrm{T}} \Delta \mathbf{w} \\ \mbox{\nabla} \mathbf{g}^{\mathrm{T}} \Delta \mathbf{w} + \mathbf{h} + \mathbf{s} \\ \mbox{\nabla} \mathbf{h}^{\mathrm{T}} \Delta \mathbf{w} + \mathbf{h} + \mathbf{s} \\ \mbox{\mu}_{i}^{\mathrm{QP}} \mathbf{s} \\ \mbox{reduce } \tau \rightarrow \epsilon \\ \mbox{end} \\ \mbox{Update} \\ \ \{\mathbf{w}, \lambda, \mu\} \leftarrow \{\mathbf{w}, \lambda, \mu\} + \Delta \{\mathbf{w}, \lambda, \mu\} \\ \end{array} $	$\mu^{\text{QP}} = 0$ $+ \mathbf{g} = 0$ $p^{\text{QP}} = 0$ $+ \mathbf{g} = 0$ $p^{\text{QP}} = \tau$ $p^$
 less linearizations more linear solves warm-start is very effective 	 more linearizations less linear solves warm-start is often ineffective