Direct Optimal Control
Lecture 11: Decomposition & Splitting

Sebastien Gros

ITK, NTNU

NTNU PhD course

Outline

Outline

1. Preliminiaries
2. Primal Decomposition
3. Dual Decomposition
4. First order method

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

Useful for e.g.

- Assessing the "price" of design parameters $\mathbf{p}$
- Decomposition methods...

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$ What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

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

Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

Useful for e.g.

- Assessing the "price" of design parameters $\mathbf{p}$
- Decomposition methods...

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable

Useful for e.g.

- Assessing the "price" of design parameters $\mathbf{p}$
- Decomposition methods...

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$

Useful for e.g.

- Assessing the "price" of design parameters $\mathbf{p}$
- Decomposition methods...

A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

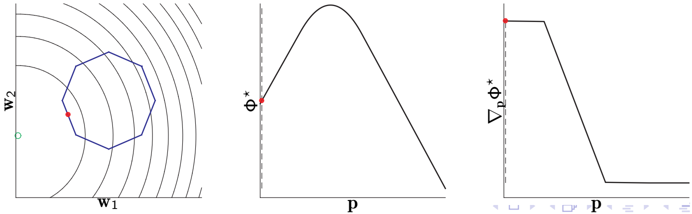
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

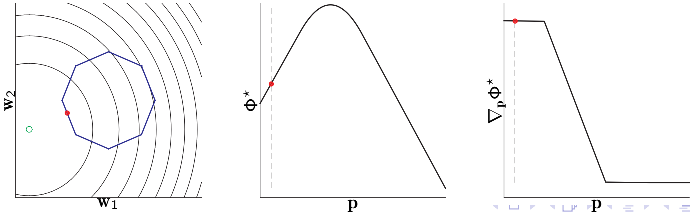
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

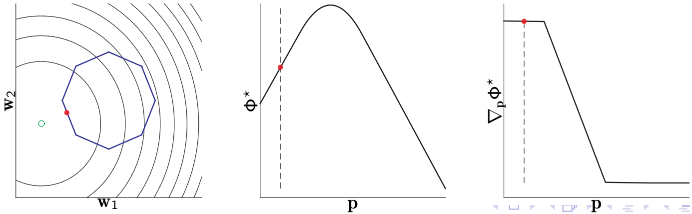
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$
\begin{aligned}
\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\
\text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\
& \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0
\end{aligned}
$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$
\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})
$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

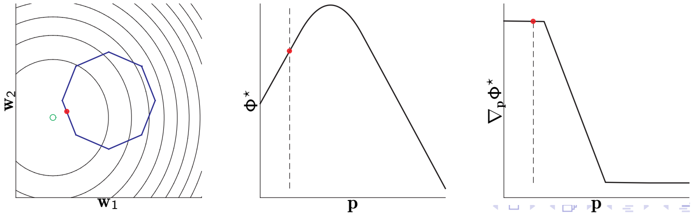
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

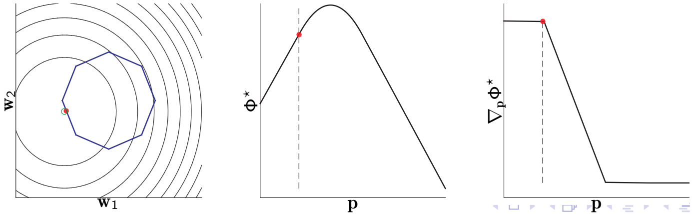
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

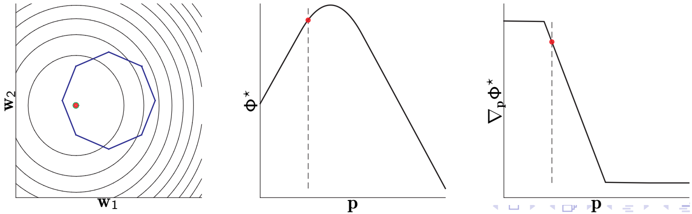
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \textcolor{red}{\arg}\min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

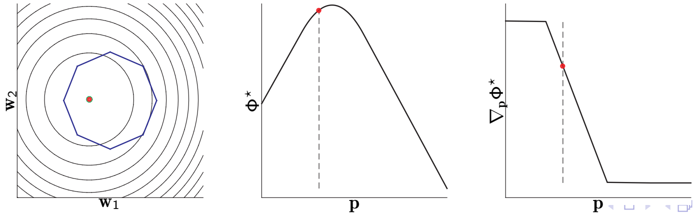
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

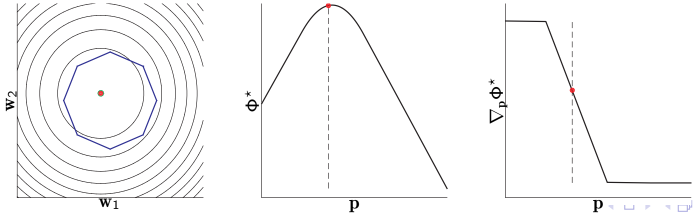
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

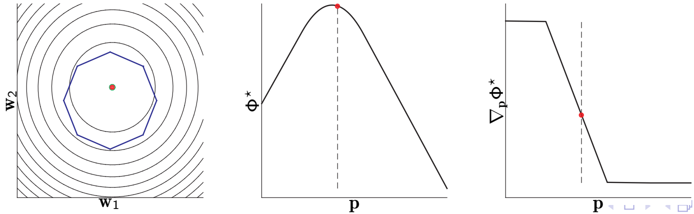
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

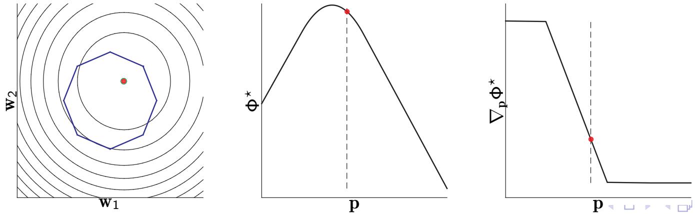
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

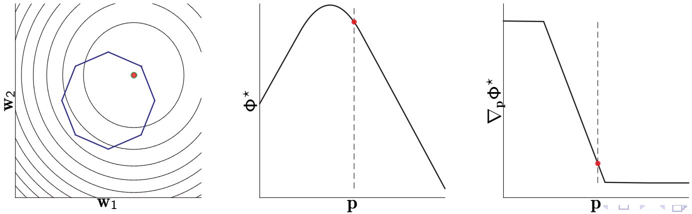
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

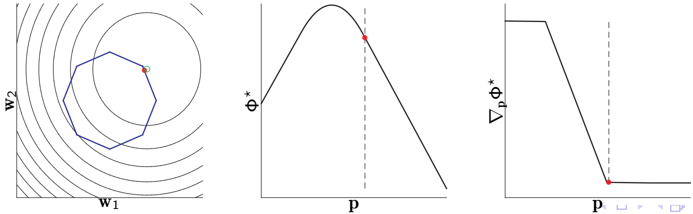
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

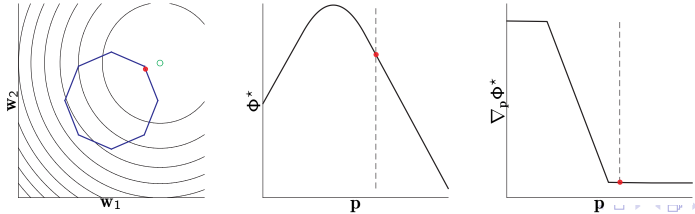
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

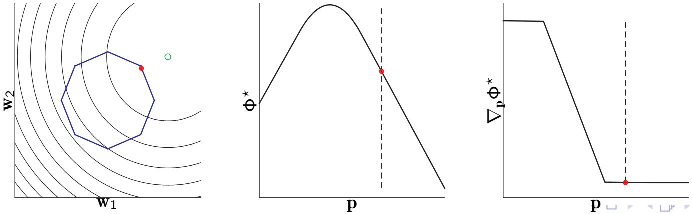
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

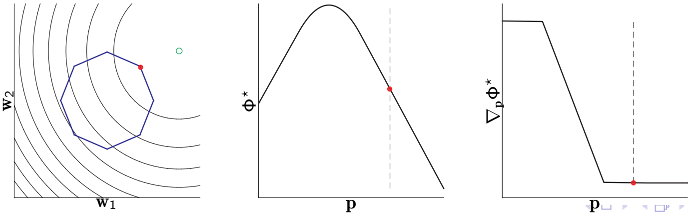
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p}) \quad \text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

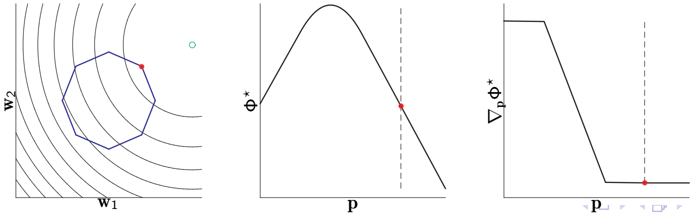
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

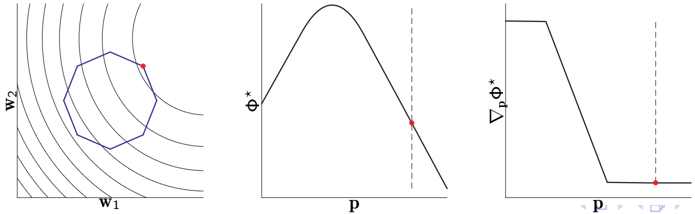
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

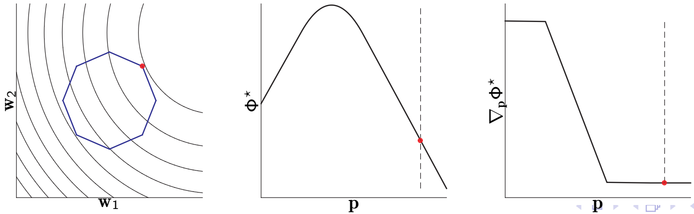
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

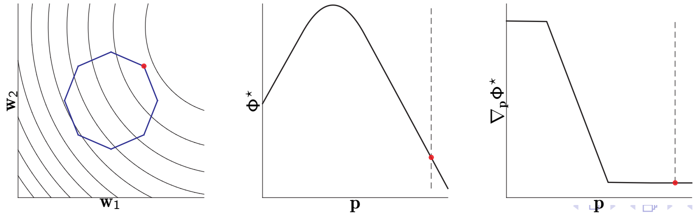
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

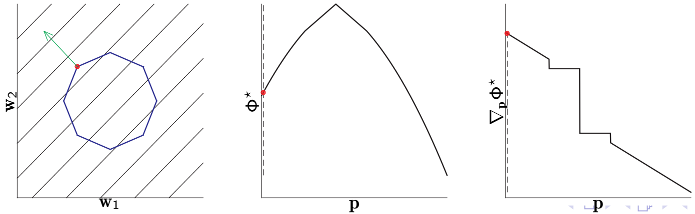
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

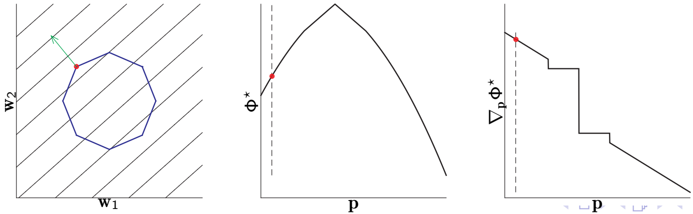
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

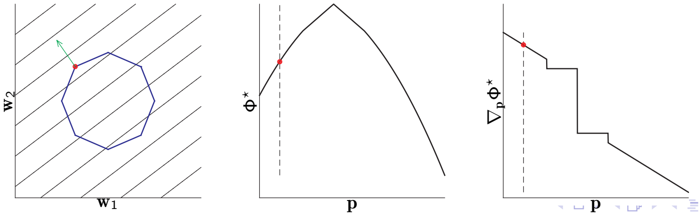
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

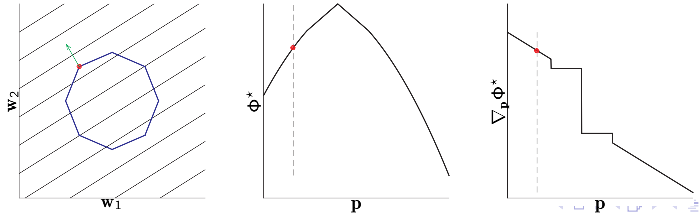
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

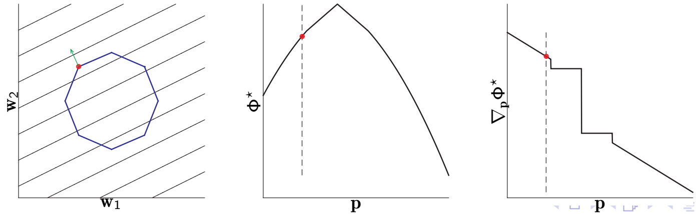
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

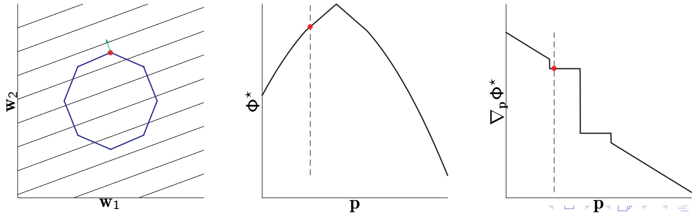
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

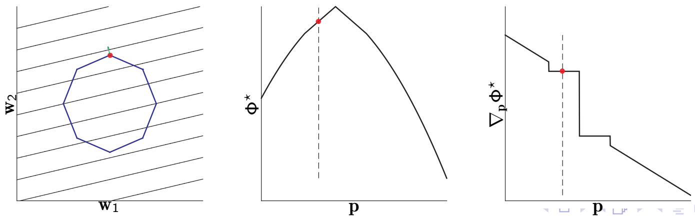
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

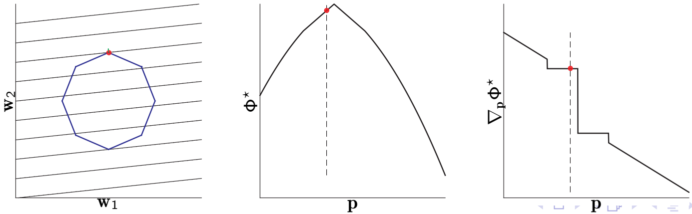
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

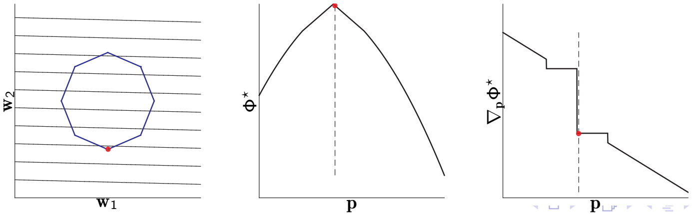
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

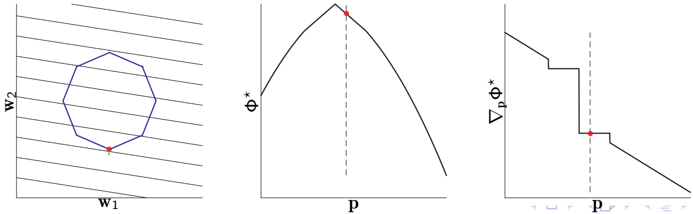
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\begin{aligned} \mathbf{w}(\mathbf{p}) = {\color{red}\arg}\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

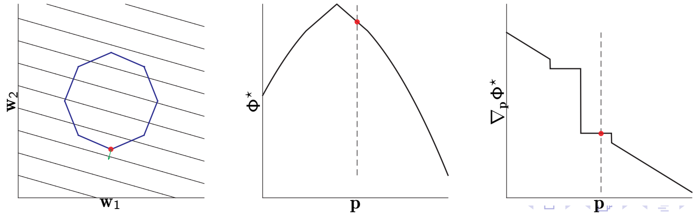
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

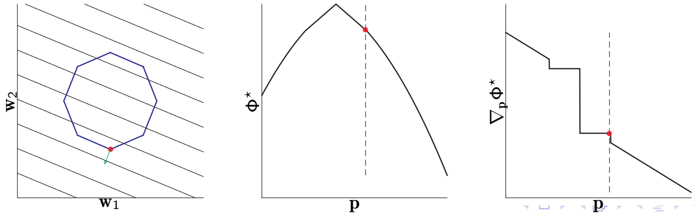
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p}) \qquad \text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

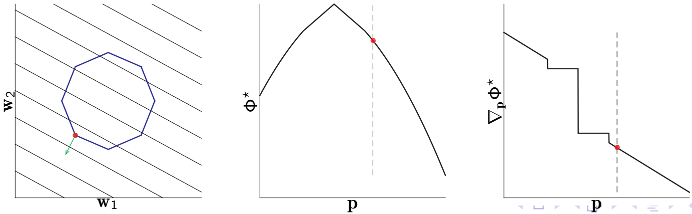
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p}) \qquad \text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

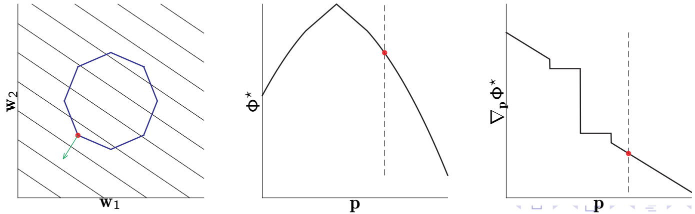
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

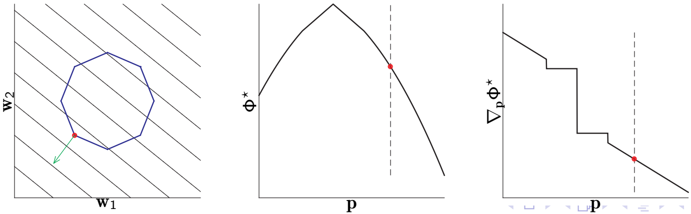
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

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

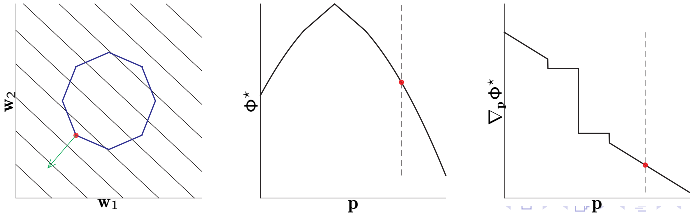
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \color{red}{\arg}\min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

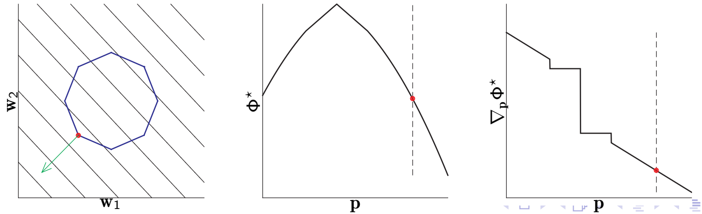
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\begin{aligned} \Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

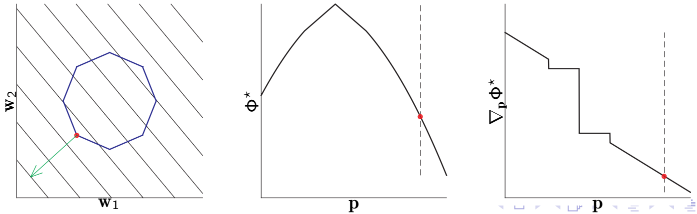
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}} \Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



A brief complement to Parametric NLPs

Parametric solution $\mathbf{w}(\mathbf{p}) \in \mathbb{R}^n$

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

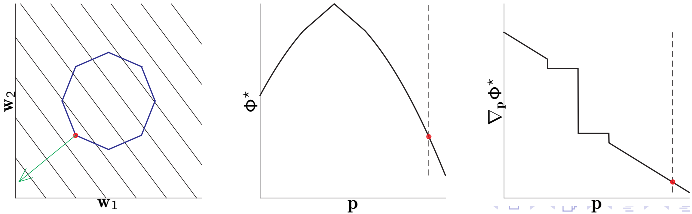
Parametric cost $\Phi^\star(\mathbf{p}) \in \mathbb{R}$

$$\Phi^\star(\mathbf{p}) = \min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

What is $\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p})$?

- If $\mathbf{w}(\mathbf{p})$ has LICQ & strict SOSC then $\Phi^\star(\mathbf{p})$ is continuously differentiable
- Sensitivity simply given by:

$$\nabla_{\mathbf{p}}\Phi^\star(\mathbf{p}) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{w}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star, \mathbf{p})$$



Outline

1. Preliminiaries
2. Primal Decomposition
3. Dual Decomposition
4. First order method

Primal Decomposition - Key idea

Original problem:

$$
\begin{aligned}
P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i(\mathbf{x}_i,\mathbf{y}) \\
\text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\
& \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0
\end{aligned}
$$

common variable $\mathbf{y}$.

Primal Decomposition - Key idea

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$.

Primal decomposition:

$$\begin{aligned} p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0 \end{aligned}$$

Primal Decomposition - Key idea

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$

Primal decomposition:

$$\begin{aligned} p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. Master problem:

$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)$ provides optimal variable $\mathbf{y}$ from which $\mathbf{x}_i^\star$ can be computed

Primal Decomposition - Key idea

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i(\mathbf{x}_i,\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\ & \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0 \end{aligned}$$

Primal decomposition:

$$\begin{aligned} p_i(\mathbf{y}) = \min_{\mathbf{x}_i} \quad & \phi_i(\mathbf{x}_i,\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\ & \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. Master problem:

$$P = \min_{\mathbf{y}} \sum_i p_i(\mathbf{y}) \text{ provides optimal variable } \mathbf{y} \text{ from which } \mathbf{x}_i^\star \text{ can be computed}$$

Lagrange functions:

$$\mathcal{L}_i(\mathbf{x}_i,\boldsymbol{\lambda}_i,\boldsymbol{\mu}_i,\mathbf{y}) = \phi_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\lambda}_i^\top \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\mu}_i^\top \mathbf{h}_i(\mathbf{x}_i,\mathbf{y})$$

Primal Decomposition - Key idea

Original problem:

$$
\begin{aligned}
P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i(\mathbf{x}_i,\mathbf{y}) \\
\text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\
& \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0
\end{aligned}
$$

Primal decomposition:

$$
\begin{aligned}
p_i(\mathbf{y}) = \min_{\mathbf{x}_i} \quad & \phi_i(\mathbf{x}_i,\mathbf{y}) \\
\text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\
& \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0
\end{aligned}
$$

common variable $\mathbf{y}$. Master problem:

$$P = \min_{\mathbf{y}} \sum_i p_i(\mathbf{y}) \text{ provides optimal variable } \mathbf{y} \text{ from which } \mathbf{x}_i^\star \text{ can be computed}$$

Lagrange functions:

$$\mathcal{L}_i(\mathbf{x}_i,\boldsymbol{\lambda}_i,\boldsymbol{\mu}_i,\mathbf{y}) = \phi_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\lambda}_i^\top \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\mu}_i^\top \mathbf{h}_i(\mathbf{x}_i,\mathbf{y})$$

- Gradient $\nabla p_i(\mathbf{y}) = \frac{\partial \mathcal{L}_i}{\partial \mathbf{y}}$ evaluated at the solution $\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\mu}_i^\star$

Primal Decomposition - Key idea

Original problem:

$$P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad \sum_i \phi_i\left(\mathbf{x}_i,\mathbf{y}\right)$$
$$\text{s.t.} \quad \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0$$
$$\mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0$$

Primal decomposition:

$$p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad \phi_i\left(\mathbf{x}_i,\mathbf{y}\right)$$
$$\text{s.t.} \quad \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0$$
$$\mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0$$

common variable $\mathbf{y}$. Master problem:

$$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right) \text{ provides optimal variable } \mathbf{y} \text{ from which } \mathbf{x}_i^\star \text{ can be computed}$$

Lagrange functions:

$$\mathcal{L}_i\left(\mathbf{x}_i,\boldsymbol{\lambda}_i,\boldsymbol{\mu}_i,\mathbf{y}\right) = \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) + \boldsymbol{\lambda}_i^\top \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) + \boldsymbol{\mu}_i^\top \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right)$$

- Gradient $\nabla p_i\left(\mathbf{y}\right) = \frac{\partial \mathcal{L}_i}{\partial \mathbf{y}}$ evaluated at the solution $\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\mu}_i^\star$
- Hessian

$$\nabla^2 p_i\left(\mathbf{y}\right) = \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}^2} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \mathbf{x}_i}\frac{\partial \mathbf{x}_i^\star}{\partial \mathbf{y}} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \boldsymbol{\lambda}_i}\frac{\partial \boldsymbol{\lambda}_i^\star}{\partial \mathbf{y}} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \boldsymbol{\mu}_i}\frac{\partial \boldsymbol{\mu}_i^\star}{\partial \mathbf{y}}$$

Primal Decomposition - Key idea

Original problem:

$$P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad \sum_i \phi_i(\mathbf{x}_i,\mathbf{y})$$
$$\text{s.t.} \quad \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0$$
$$\mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0$$

Primal decomposition:

$$p_i(\mathbf{y}) = \min_{\mathbf{x}_i} \quad \phi_i(\mathbf{x}_i,\mathbf{y})$$
$$\text{s.t.} \quad \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0$$
$$\mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0$$

common variable $\mathbf{y}$. Master problem:

$$P = \min_{\mathbf{y}} \sum_i p_i(\mathbf{y}) \text{ provides optimal variable } \mathbf{y} \text{ from which } \mathbf{x}_i^\star \text{ can be computed}$$

Lagrange functions:

$$\mathcal{L}_i(\mathbf{x}_i,\boldsymbol{\lambda}_i,\boldsymbol{\mu}_i,\mathbf{y}) = \phi_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\lambda}_i^\top \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) + \boldsymbol{\mu}_i^\top \mathbf{h}_i(\mathbf{x}_i,\mathbf{y})$$

- Gradient $\nabla p_i(\mathbf{y}) = \frac{\partial \mathcal{L}_i}{\partial \mathbf{y}}$ evaluated at the solution $\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\mu}_i^\star$
- Hessian

$$\nabla^2 p_i(\mathbf{y}) = \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}^2} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \mathbf{x}_i}\frac{\partial \mathbf{x}_i^\star}{\partial \mathbf{y}} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \boldsymbol{\lambda}_i}\frac{\partial \boldsymbol{\lambda}_i^\star}{\partial \mathbf{y}} + \frac{\partial^2 \mathcal{L}_i}{\partial \mathbf{y}\partial \boldsymbol{\mu}_i}\frac{\partial \boldsymbol{\mu}_i^\star}{\partial \mathbf{y}}$$

where $\frac{\partial \mathbf{x}_i^\star}{\partial \mathbf{y}}$, $\frac{\partial \boldsymbol{\lambda}_i^\star}{\partial \mathbf{y}}$, $\frac{\partial \boldsymbol{\mu}_i^\star}{\partial \mathbf{y}}$ is obtained via parametric optimization principles.

Primal Decomposition - Deploying the decomposition

Original problem:

$$
\begin{aligned}
P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i(\mathbf{x}_i,\mathbf{y}) \\
\text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\
& \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0
\end{aligned}
$$

common variable $\mathbf{y}$. "Master" problem:

$$
P = \min_{\mathbf{y}} \sum_i p_i(\mathbf{y})
$$

Primal decomposition:

$$
\begin{aligned}
p_i(\mathbf{y}) = \min_{\mathbf{x}_i} \quad & \phi_i(\mathbf{x}_i,\mathbf{y}) \\
\text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\
& \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0
\end{aligned}
$$

Primal Decomposition - Deploying the decomposition

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0 \end{aligned}$$

Primal decomposition:

$$\begin{aligned} p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. "Master" problem:

$$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)$$

- Evaluation of $p_i\left(\mathbf{y}\right)$ requires solving a classic optimization problem
- Evaluation of the sensitivities $\frac{\partial \mathbf{x}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\lambda}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\mu}^\star}{\partial \mathbf{y}}$ can be done cheaply using results from parametric NLPs (c.f. corresponding lecture !)
- From $\frac{\partial \mathbf{x}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\lambda}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\mu}^\star}{\partial \mathbf{y}}$ one can directly assemble $\nabla p_i\left(\mathbf{y}\right)$ and $\nabla^2 p_i\left(\mathbf{y}\right)$

Primal Decomposition - Deploying the decomposition

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i(\mathbf{x}_i,\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\ & \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0 \end{aligned}$$

Primal decomposition:

$$\begin{aligned} p_i(\mathbf{y}) = \min_{\mathbf{x}_i} \quad & \phi_i(\mathbf{x}_i,\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{g}_i(\mathbf{x}_i,\mathbf{y}) = 0 \\ & \mathbf{h}_i(\mathbf{x}_i,\mathbf{y}) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. "Master" problem:

$$P = \min_{\mathbf{y}} \sum_i p_i(\mathbf{y})$$

- Evaluation of $p_i(\mathbf{y})$ requires solving a classic optimization problem
- Evaluation of the sensitivities $\frac{\partial \mathbf{x}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\lambda}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\mu}^\star}{\partial \mathbf{y}}$ can be done cheaply using results from parametric NLPs (c.f. corresponding lecture !)
- From $\frac{\partial \mathbf{x}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\lambda}^\star}{\partial \mathbf{y}}, \frac{\partial \boldsymbol{\mu}^\star}{\partial \mathbf{y}}$ one can directly assemble $\nabla p_i(\mathbf{y})$ and $\nabla^2 p_i(\mathbf{y})$
- Steps on the common variables $\mathbf{y}$:
  - Gradient step: $\mathbf{y}^+ = \mathbf{y} - \frac{1}{L}\sum_i \nabla p_i(\mathbf{y})$
  - Newton step: $\mathbf{y}^+ = \mathbf{y} - \left(\sum_i \nabla^2 p_i(\mathbf{y})\right)^{-1} \sum_i \nabla p_i(\mathbf{y})$

Primal Decomposition - Deploying the decomposition

Original problem:

$$\begin{aligned} P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$

Primal decomposition:

$$\begin{aligned} p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. "Master" problem:

$$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)$$

- Steps on the common variables $\mathbf{y}$:
  - Gradient step: $\mathbf{y}^+ = \mathbf{y} - \frac{1}{L}\sum_i \nabla p_i\left(\mathbf{y}\right)$
  - Newton step: $\mathbf{y}^+ = \mathbf{y} - \left(\sum_i \nabla^2 p_i\left(\mathbf{y}\right)\right)^{-1} \sum_i \nabla p_i\left(\mathbf{y}\right)$
- Carefull: functions $p_i\left(\mathbf{y}\right)$ are not everywhere *twice* continuously differentiable !
- Newton steps may require careful backtracking, as the validity of the quadratic model can be very local !

Primal Decomposition - Deploying the decomposition

Original problem:

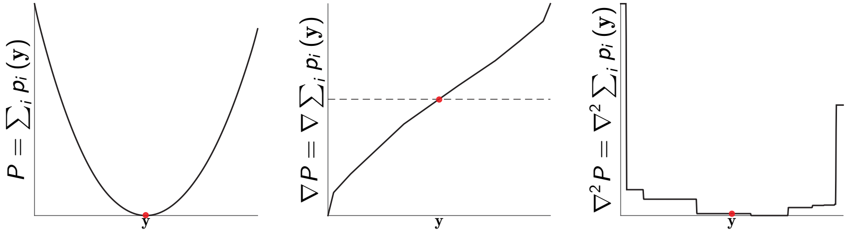
$$
\begin{aligned}
P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\
\text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\
& \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0
\end{aligned}
$$

Primal decomposition:

$$
\begin{aligned}
p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i,\mathbf{y}\right) \\
\text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0 \\
& \mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0
\end{aligned}
$$

common variable $\mathbf{y}$. "Master" problem:

$$
P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)
$$



Primal Decomposition - Deploying the decomposition

Original problem:

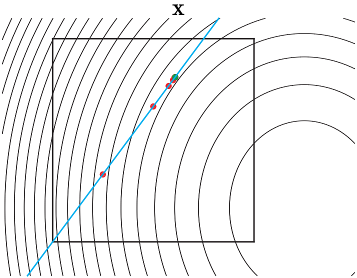
$$\begin{aligned} P = \min_{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}} \quad & \sum_i \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$

common variable $\mathbf{y}$. "Master" problem:

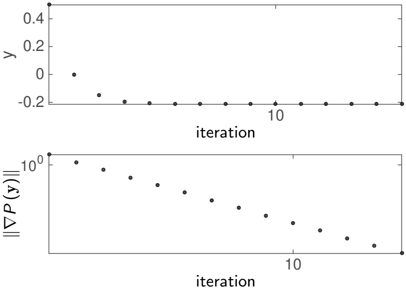
$$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)$$

Primal decomposition:

$$\begin{aligned} p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad & \phi_i\left(\mathbf{x}_i, \mathbf{y}\right) \\ \text{s.t.} \quad & \mathbf{g}_i\left(\mathbf{x}_i, \mathbf{y}\right) = 0 \\ & \mathbf{h}_i\left(\mathbf{x}_i, \mathbf{y}\right) \leq 0 \end{aligned}$$



First-order steps on $\mathbf{y}$



Primal Decomposition - Deploying the decomposition

Original problem:

$$P = \min_{\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}} \quad \sum_i \phi_i\left(\mathbf{x}_i,\mathbf{y}\right)$$
$$\text{s.t.} \quad \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0$$
$$\mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0$$

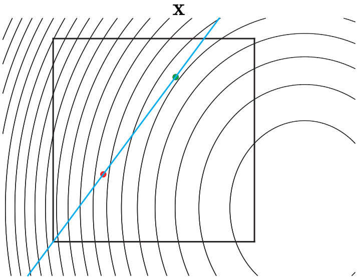
Primal decomposition:

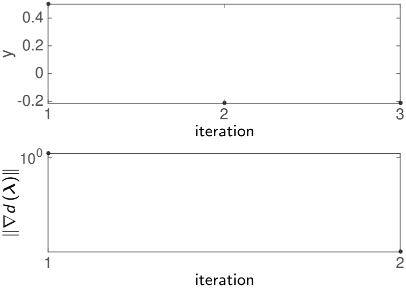
$$p_i\left(\mathbf{y}\right) = \min_{\mathbf{x}_i} \quad \phi_i\left(\mathbf{x}_i,\mathbf{y}\right)$$
$$\text{s.t.} \quad \mathbf{g}_i\left(\mathbf{x}_i,\mathbf{y}\right) = 0$$
$$\mathbf{h}_i\left(\mathbf{x}_i,\mathbf{y}\right) \leq 0$$

common variable $\mathbf{y}$. "Master" problem:

$$P = \min_{\mathbf{y}} \sum_i p_i\left(\mathbf{y}\right)$$

Second-order steps on $\mathbf{y}$





Primal Decomposition - Example of application

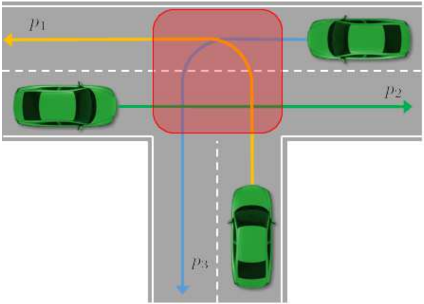
Coordination of cars approaching an intersection

Given ordering S, master problem:

- in and out times $T_j^{\mathrm{in}}, \ T_j^{\mathrm{out}}$ of each vehicle j

$$
\begin{array}{lll}
\min\limits_{T} & \sum_{j=1}^{M} \Phi_j(T_j) & \text{Cost for each car} \\
\text{s.t.} & T_j \in \mathbb{T}_j & \text{Feasible in \& out times} \\
 & T_{S_{i+1}}^{\mathrm{in}} \geq T_{S_i}^{\mathrm{out}} & \text{No collision (coupling)}
\end{array}
$$

Times T are assigned to the car by the master problem



Coordination of cars approaching an intersection

Given ordering S, master problem:

- in and out times T_j^{in}, T_j^{out} of each vehicle j

$$\begin{aligned} \min_{T} \quad & \textstyle\sum_{j=1}^{M} \Phi_j(T_j) && \text{Cost for each car} \\ \text{s.t.} \quad & T_j \in \mathbb{T}_j && \text{Feasible in \& out times} \\ & T^{\mathrm{in}}_{S_{i+1}} \geq T^{\mathrm{out}}_{S_i} && \text{No collision (coupling)} \end{aligned}$$

Times T are assigned to the car by the master problem

Primal decomposition:

- MPC for each car:

$$\begin{aligned} \Phi_j(T_j) := \min_{x,u} \quad & V_{\mathrm{f}}(x_N^j) + \sum_{k=0}^{N-1} \ell(x_k^j, u_k^j) \\ \text{s.t.} \quad & x_0^j = \hat{x}_0^j \\ & x_{k+1}^j = A_j x_k^j + B_j u_k^j \\ & C_j x_k^j + D_j u_k^j \geq 0 \\ & p(T_j^{\mathrm{in}}, x^j, u^j) = p_j^{\mathrm{in}} \\ & p(T_j^{\mathrm{out}}, x^j, u^j) = p_j^{\mathrm{out}} \end{aligned}$$

$$\mathbb{T}_j := \mathrm{dom}(\Phi_j(T_j))$$

Outline

Dual Decomposition - Key idea

Consider

$$\min_{\mathbf{x}} \quad \phi(\mathbf{x}) \tag{1a}$$

$$\text{s.t.} \quad A\mathbf{x} + \mathbf{b} = 0 \tag{1b}$$

$$C\mathbf{x} \leq \mathbf{d} \tag{1c}$$

Dual Decomposition - Key idea

Consider

$$\min_{\mathbf{x}} \quad \phi(\mathbf{x}) \tag{1a}$$
$$\text{s.t.} \quad A\mathbf{x} + \mathbf{b} = 0 \tag{1b}$$
$$C\mathbf{x} \leq \mathbf{d} \tag{1c}$$

And define the *partial* Lagrange function:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$$

Dual Decomposition - Key idea

Consider

$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) && \text{(1a)} \\ \text{s.t.} \quad & A\mathbf{x} + \mathbf{b} = 0 && \text{(1b)} \\ & C\mathbf{x} \le \mathbf{d} && \text{(1c)} \end{aligned}$$

And define the *partial* Lagrange function:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$$

We can then define the dual function:

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) && \text{(2a)} \\ \text{s.t.} \quad & C\mathbf{x} \le \mathbf{d} && \text{(2b)} \end{aligned}$$

Dual Decomposition - Key idea

Consider

$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) && \text{(1a)} \\ \text{s.t.} \quad & A\mathbf{x} + \mathbf{b} = 0 && \text{(1b)} \\ & C\mathbf{x} \leq \mathbf{d} && \text{(1c)} \end{aligned}$$

And define the *partial* Lagrange function:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$$

We can then define the dual function:

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) && \text{(2a)} \\ \text{s.t.} \quad & C\mathbf{x} \leq \mathbf{d} && \text{(2b)} \end{aligned}$$

If (1) has strong duality, then solution can be computed via the **dual problem**:

$$\min_{\boldsymbol{\lambda}} \quad d(\boldsymbol{\lambda})$$

and the minimizer $\mathbf{x}$ of the dual problem (2) is the primal solution of (1)

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

Dualization of the equality constraints

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \mathrm{s.t.} \quad & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

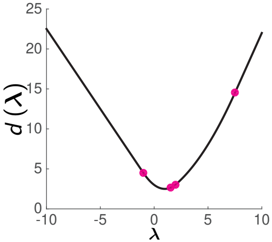
$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

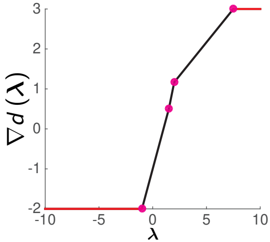
Dualization of the equality constraints

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

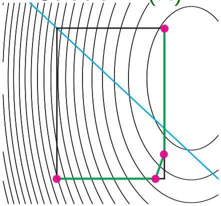
where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Behavior of dual function:





Solution $\mathbf{x}(\boldsymbol{\lambda})$



Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

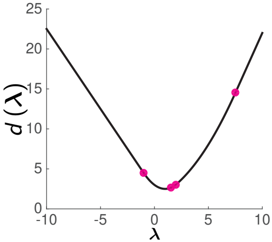
$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

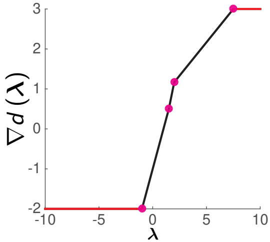
Dualization of the equality constraints

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

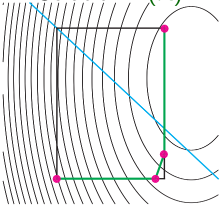
where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Behavior of dual function:





Solution $\mathbf{x}(\boldsymbol{\lambda})$



- $d(\boldsymbol{\lambda})$ convex (if original problem is)

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

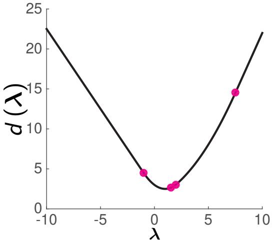
$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

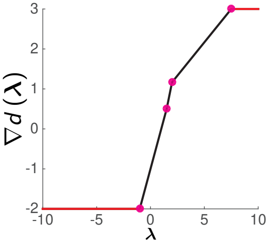
Dualization of the equality constraints

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

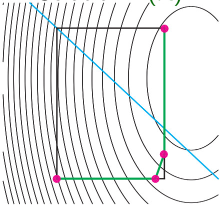
where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Behavior of dual function:





Solution $\mathbf{x}(\boldsymbol{\lambda})$



- $d(\boldsymbol{\lambda})$ convex (if original problem is)
- $\nabla d(\boldsymbol{\lambda})$ continuous, non-differentiable when active set changes

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

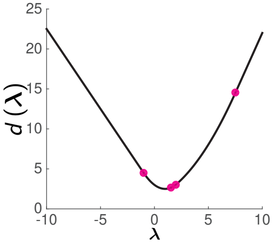
$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

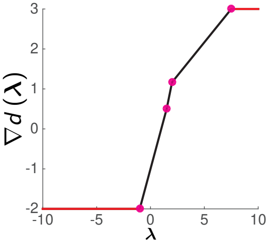
Dualization of the equality constraints

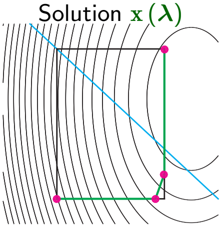
$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Behavior of dual function:







- $d(\boldsymbol{\lambda})$ convex (if original problem is)
- $\nabla d(\boldsymbol{\lambda})$ continuous, non-differentiable when active set changes
- $\nabla^2 d(\boldsymbol{\lambda})$ loses rank if dualised constraints + remaining constraints fail LICQ

Dual Decomposition - Some properties of the dual function

Example with $\mathbf{x} \in \mathbb{R}^2$

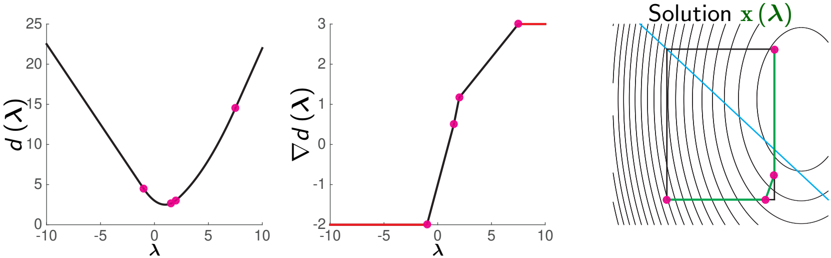
$$\begin{aligned} \min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\ & A\mathbf{x} + \mathbf{b} = 0 \\ & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

Dualization of the equality constraints

$$\begin{aligned} d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \phi(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} + \mathbf{b})$

Behavior of dual function:



- $d(\boldsymbol{\lambda})$ convex (if original problem is)
- $\nabla d(\boldsymbol{\lambda})$ continuous, non-differentiable when active set changes
- $\nabla^2 d(\boldsymbol{\lambda})$ loses rank if dualised constraints + remaining constraints fail LICQ
- $\nabla d(\boldsymbol{\lambda}) = A\mathbf{x} + \mathbf{b}$, i.e. residual of dualized constraints provides dual gradient !

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$
\begin{aligned}
\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) \\
\text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, ..., N \\
& C_i \mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, ..., N \\
& \sum_{i=1}^{N} D_i \mathbf{x}_i = 0
\end{aligned}
$$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$
\begin{aligned}
\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) \\
\text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, ..., N \\
& C_i \mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, ..., N \\
& \sum_{i=1}^{N} D_i \mathbf{x}_i = 0
\end{aligned}
$$

"Dualized" coupling constraints:

$$
\begin{aligned}
\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i \mathbf{x}_i \right) \\
\text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\
& C_i \mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$\begin{aligned} \min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, ..., N \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, ..., N \\ & \sum_{i=1}^{N} D_i \mathbf{x}_i = 0 \end{aligned}$$

"Dualized" coupling constraints:

$$\begin{aligned} \min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i \mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

Dual function reads as:

$$\begin{aligned} d\left(\boldsymbol{\lambda}\right) = - \min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i \mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$
\begin{aligned}
\min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) \\
\text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, \ldots, N \\
& C_i\mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, \ldots, N \\
& \sum_{i=1}^{N} D_i\mathbf{x}_i = 0
\end{aligned}
$$

"Dualized" coupling constraints:

$$
\begin{aligned}
\min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i\mathbf{x}_i\right) \\
\text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\
& C_i\mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

Dual problem reads as:

$$
\begin{aligned}
d\left(\boldsymbol{\lambda}\right) = -\min_{\mathbf{x}_1,\ldots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \mathcal{L}_i\left(\mathbf{x}_i, \boldsymbol{\lambda}\right) \\
\text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\
& C_i\mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

where $\mathcal{L}_i\left(\mathbf{x}_i, \boldsymbol{\lambda}\right) = \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top D_i\mathbf{x}_i$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$\begin{aligned}\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, ..., N \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, ..., N \\ & \sum_{i=1}^{N} D_i\mathbf{x}_i = 0\end{aligned}$$

"Dualized" coupling constraints:

$$\begin{aligned}\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i\mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i\end{aligned}$$

Dual problem reads as:

$$\begin{aligned}d(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i\end{aligned}$$

where $\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda}) = \phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i\mathbf{x}_i$

Define:

$$\begin{aligned}d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i\end{aligned}$$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$\begin{aligned} \min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1, ..., N \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1, ..., N \\ & \sum_{i=1}^{N} D_i \mathbf{x}_i = 0 \end{aligned}$$

The **dual problem** then reads as:

$$d\left(\boldsymbol{\lambda}\right) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i\left(\boldsymbol{\lambda}\right)$$

"Dualized" coupling constraints:

$$\begin{aligned} \min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i\left(\mathbf{x}_i\right) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i \mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

Define:

$$\begin{aligned} d_i\left(\boldsymbol{\lambda}\right) = -\min_{\mathbf{x}_i} \quad & \mathcal{L}_i\left(\mathbf{x}_i, \boldsymbol{\lambda}\right) \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

Dual Decomposition - Decomposable Problems

Consider a problem in the form:

$$\begin{aligned}\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i, \quad \forall\, i = 1,\dots,N \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i, \quad \forall\, i = 1,\dots,N \\ & \sum_{i=1}^{N} D_i\mathbf{x}_i = 0\end{aligned}$$

"Dualized" coupling constraints:

$$\begin{aligned}\min_{\mathbf{x}_1,\dots,\mathbf{x}_N} \quad & \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top \left(\sum_{i=1}^{N} D_i\mathbf{x}_i\right) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i\end{aligned}$$

Define:

$$\begin{aligned}d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\ & C_i\mathbf{x}_i \leq \mathbf{d}_i\end{aligned}$$

The **dual problem** then reads as:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Large, coupled problem turned into N small, decoupled subproblems $d_i(\boldsymbol{\lambda})$
- The dual problem $d(\boldsymbol{\lambda})$ and associated multipliers $\boldsymbol{\lambda}$ act as "coordinator" between the subproblems
- Dual variable $\boldsymbol{\lambda}$ modifies the gradients in the "local" problems $d_{1,\dots,N}$.

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$\begin{aligned} d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i \mathbf{x}_i} \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$
\begin{aligned}
d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i \mathbf{x}_i} \\
\text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\
& C_i \mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$
\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})
$$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$\begin{aligned} d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i \mathbf{x}_i} \\ \text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\ & C_i \mathbf{x}_i \leq \mathbf{d}_i \end{aligned}$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i \frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$

$$\text{s.t.} \quad A_i \mathbf{x}_i = \mathbf{b}_i$$

$$C_i \mathbf{x}_i \leq \mathbf{d}_i$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i \frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

- Dual problem:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Evaluating $d_i(\boldsymbol{\lambda})$ requires solving a classic convex optimization problem

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i\mathbf{x}_i}$$
$$\text{s.t.} \quad A_i\mathbf{x}_i = \mathbf{b}_i$$
$$C_i\mathbf{x}_i \leq \mathbf{d}_i$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i\mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i\frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

- Dual problem:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Evaluating $d_i(\boldsymbol{\lambda})$ requires solving a classic convex optimization problem
- Upon having evaluation of all $d_i(\boldsymbol{\lambda})$, one can take a step on the multipliers $\boldsymbol{\lambda}$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$
$$\text{s.t.} \quad A_i \mathbf{x}_i = \mathbf{b}_i$$
$$C_i \mathbf{x}_i \leq \mathbf{d}_i$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i \frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

- Dual problem:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Evaluating $d_i(\boldsymbol{\lambda})$ requires solving a classic convex optimization problem
- Upon having evaluation of all $d_i(\boldsymbol{\lambda})$, one can take a step on the multipliers $\boldsymbol{\lambda}$
- Gradient step:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{L} \sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$
$$\text{s.t.} \quad A_i \mathbf{x}_i = \mathbf{b}_i$$
$$C_i \mathbf{x}_i \leq \mathbf{d}_i$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i \frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

- Dual problem:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Evaluating $d_i(\boldsymbol{\lambda})$ requires solving a classic convex optimization problem
- Upon having evaluation of all $d_i(\boldsymbol{\lambda})$, one can take a step on the multipliers $\boldsymbol{\lambda}$
- Gradient step:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{L}\sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$

- Newton step:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \left[\sum_{i=1}^{N} \nabla^2 d_i(\boldsymbol{\lambda})\right]^{-1} \sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$

Dual Decomposition - Deploying the Dual Decomposition

Decomposed subproblems

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$

$$\text{s.t.} \quad A_i \mathbf{x}_i = \mathbf{b}_i$$

$$C_i \mathbf{x}_i \le \mathbf{d}_i$$

yields $\mathbf{x}_i^\star(\boldsymbol{\lambda})$

- Gradient:

$$\nabla d_i(\boldsymbol{\lambda}) = -D_i \mathbf{x}_i^\star(\boldsymbol{\lambda})$$

- Hessian:

$$\nabla^2 d_i(\boldsymbol{\lambda}) = -D_i \frac{\partial \mathbf{x}_i^\star(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}$$

- Dual problem:

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \sum_{i=1}^{N} d_i(\boldsymbol{\lambda})$$

- Evaluating $d_i(\boldsymbol{\lambda})$ requires solving a classic convex optimization problem
- Upon having evaluation of all $d_i(\boldsymbol{\lambda})$, one can take a step on the multipliers $\boldsymbol{\lambda}$
- Gradient step:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{L}\sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$

- Newton step:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \left[\sum_{i=1}^{N} \nabla^2 d_i(\boldsymbol{\lambda})\right]^{-1} \sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$

- Dual function $d(\boldsymbol{\lambda})$ is not everywhere twice differentiable (change of Active Set). Newton steps may require careful backtracking !!

Dual Decomposition - Example

Local problems:

$$
\begin{aligned}
d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i \mathbf{x}_i} \\
\text{s.t.} \quad & A_i \mathbf{x}_i = \mathbf{b}_i \\
& C_i \mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

Dual Decomposition - Example

Local problems:

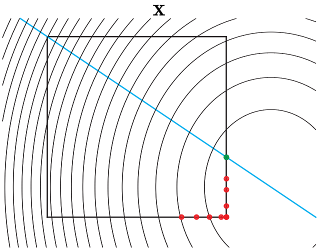
$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$

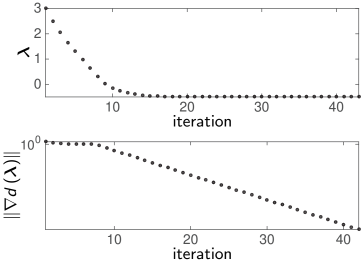
$$\text{s.t.} \quad A_i \mathbf{x}_i = \mathbf{b}_i$$

$$C_i \mathbf{x}_i \leq \mathbf{d}_i$$

First-order dual steps

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{L} \sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$





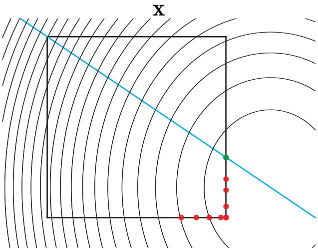
Dual Decomposition - Example

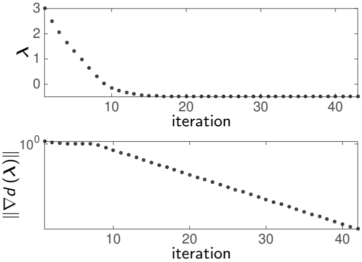
Local problems:

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i\mathbf{x}_i}$$
$$\text{s.t.} \quad A_i\mathbf{x}_i = \mathbf{b}_i$$
$$C_i\mathbf{x}_i \leq \mathbf{d}_i$$

First-order dual steps

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{L}\sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})$$





- Linear convergence
- Dual step inexpensive, using $\nabla d_i(\boldsymbol{\lambda}) = -D_i\mathbf{x}_i^\star$ and no linear algebra

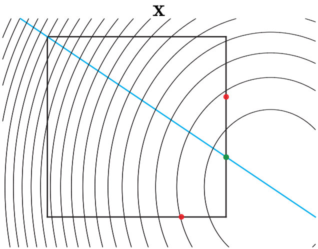
Dual Decomposition - Example

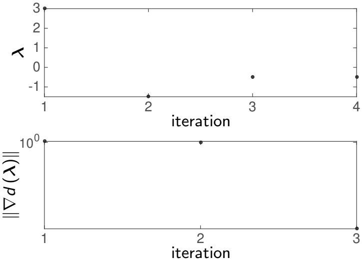
Local problems:

$$d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i \mathbf{x}_i}$$
$$\text{s.t.} \quad A_i\mathbf{x}_i = \mathbf{b}_i$$
$$C_i\mathbf{x}_i \leq \mathbf{d}_i$$

Second-order (full) dual steps

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \left[\sum_{i=1}^N \nabla^2 d_i(\boldsymbol{\lambda})\right]^{-1} \sum_{i=1}^N \nabla d_i(\boldsymbol{\lambda})$$





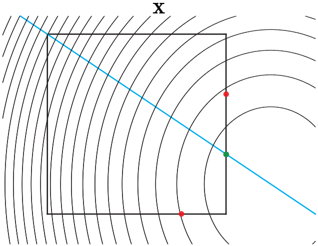
Dual Decomposition - Example

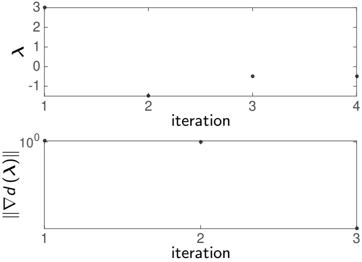
Local problems:

$$
\begin{aligned}
d_i(\boldsymbol{\lambda}) = -\min_{\mathbf{x}_i} \quad & \overbrace{\mathcal{L}_i(\mathbf{x}_i, \boldsymbol{\lambda})}^{\phi_i(\mathbf{x}_i)+\boldsymbol{\lambda}^\top D_i\mathbf{x}_i} \\
\text{s.t.} \quad & A_i\mathbf{x}_i = \mathbf{b}_i \\
& C_i\mathbf{x}_i \leq \mathbf{d}_i
\end{aligned}
$$

Second-order (full) dual steps

$$
\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \left[\sum_{i=1}^{N} \nabla^2 d_i(\boldsymbol{\lambda})\right]^{-1} \sum_{i=1}^{N} \nabla d_i(\boldsymbol{\lambda})
$$

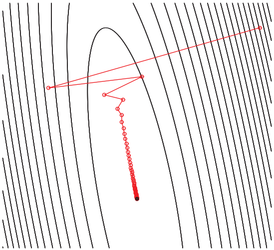


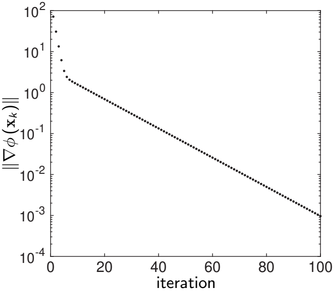


- Once correct active set identified, convergence in one step (local QPs)
- Dual Hessian can be rank deficient !! (LICQ deficiency for some $\boldsymbol{\lambda}$)
- Non-smooth Newton requires careful backtracking (not done in this example)

Outline

A bit of background - Gradient method





Problem:

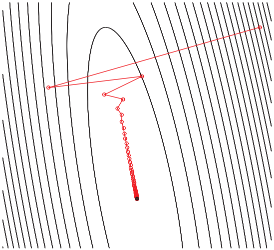
$$\min_{\mathbf{x}} \quad \phi(\mathbf{x})$$

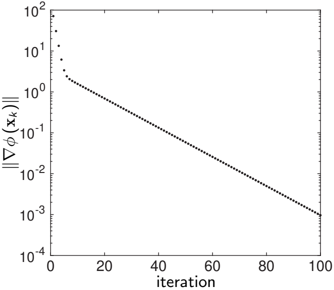
Basic gradient step: start with $\mathbf{x}_0$, do

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)$$

where L is a Lipschitz constant for $\nabla\phi$.

A bit of background - Gradient method



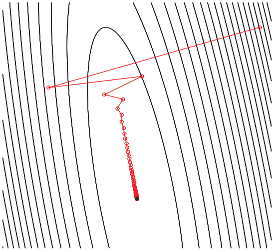


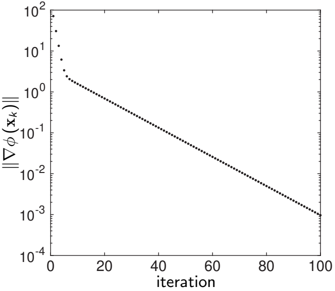
Basic gradient step: start with $\mathbf{x}_0$, do

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{L}\nabla\phi\left(\mathbf{x}_k\right)$$

where L is the Lipschitz constant of $\nabla\phi$.

A bit of background - Gradient method





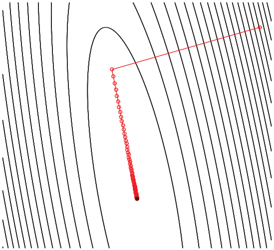
Basic gradient step: start with $\mathbf{x}_0$, do

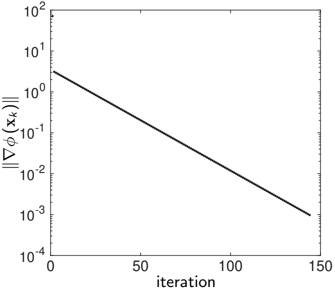
$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{L}\nabla\phi\left(\mathbf{x}_k\right)$$

where L is the Lipschitz constant of $\nabla\phi$. Observe that a Newton step reads as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \nabla^2\phi\left(\mathbf{x}_k\right)^{-1}\nabla\phi\left(\mathbf{x}_k\right)$$

A bit of background - Gradient method





Basic gradient step: start with $\mathbf{x}_0$, do

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{L}\nabla\phi\left(\mathbf{x}_k\right)$$

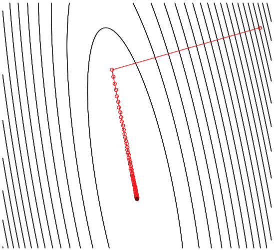
where L is the Lipschitz constant of $\nabla\phi$. Observe that a Newton step reads as:

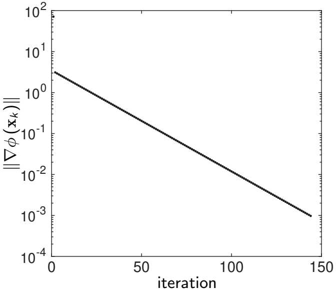
$$\mathbf{x}_{k+1} = \mathbf{x}_k - \nabla^2\phi\left(\mathbf{x}_k\right)^{-1}\nabla\phi\left(\mathbf{x}_k\right)$$

Step-size $\frac{1}{L}$ serves as a crude approximation of $\nabla^2\phi^{-1}$. Indeed, upper bound:

$$\nabla^2\phi\left(\mathbf{x}_k\right) \leq L\mathbb{I} \quad \text{ensures convergence of gradient step}$$

A bit of background - Gradient method





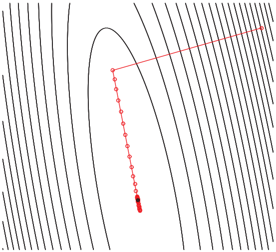
Fast gradient method, $\mathbf{x}_0 = \mathbf{y}_0$, does:

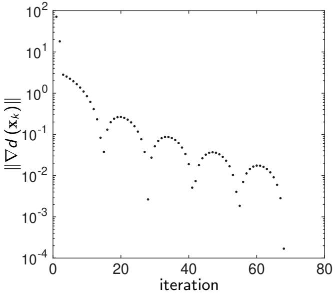
$$\mathbf{x}_k = \mathbf{y}_{k-1} - \frac{1}{L}\nabla\phi\left(\mathbf{y}_{k-1}\right)$$
$$\mathbf{y}_k = \mathbf{x}_k + \frac{k-1}{k+2}\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right)$$

"Heavy-ball" interpretation

- Introduce some "inertia" in the system
- If the gradient steps are consistently in the same direction, then larger steps are taken

A bit of background - Gradient method





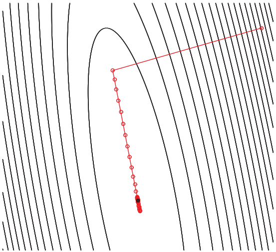
Fast gradient method, $\mathbf{x}_0 = \mathbf{y}_0$, does:

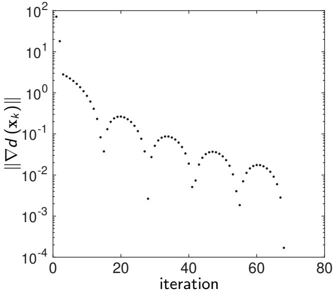
$$\mathbf{x}_k = \mathbf{y}_{k-1} - \frac{1}{L}\nabla\phi\left(\mathbf{y}_{k-1}\right)$$
$$\mathbf{y}_k = \mathbf{x}_k + \frac{k-1}{k+2}\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right)$$

"Heavy-ball" interpretation

- Introduce some "inertia" in the system
- If the gradient steps are consistently in the same direction, then larger steps are taken

A bit of background - Gradient method





What about a constrained problem ?

$$\min_{\mathbf{x} \in X} \quad \phi(\mathbf{x})$$

$\rightarrow$ **Proximal gradient method**

Proximal gradient method

Constrained problem:

$$\min_{\mathbf{x} \in X} \quad \phi(\mathbf{x})$$

Proximal gradient method

Constrained problem:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_{\mathcal{X}}(\mathbf{y}) = \arg\min_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x} - \mathbf{y}\|^2$$

Proximal gradient method

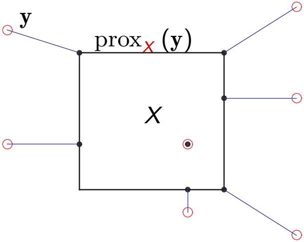
Constrained problem:

$$\min_{\mathbf{x} \in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\text{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|^2$$

Prox operator when X is a box



Proximal gradient method

Constrained problem:

$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

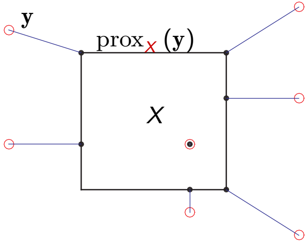
Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$

Prox operator when X is a box



Proximal gradient method

Constrained problem:

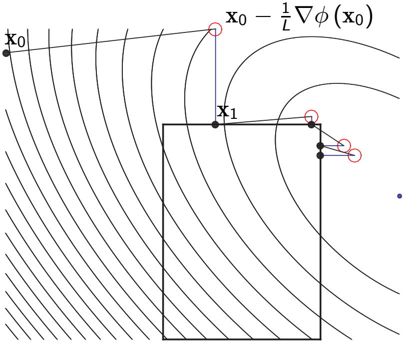
$$\min_{\mathbf{x} \in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Proximal gradient method

Constrained problem:

$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

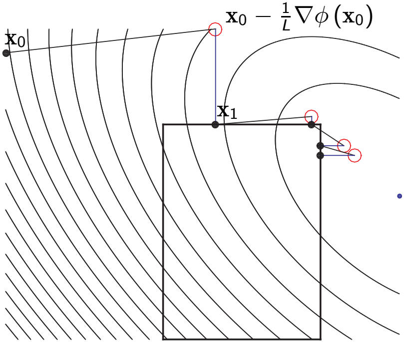
Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$

Fast proximal gradient is a direct extension



Proximal gradient method

Constrained problem:

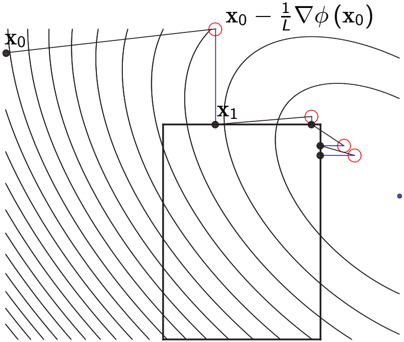
$$\min_{\mathbf{x}\in \color{red}{X}} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_{\color{red}{X}}(\mathbf{y}) = \arg\min_{\mathbf{x}\in \color{red}{X}} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_{\color{red}{X}}\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Special case: **clipping method**. Consider

$$\begin{aligned}\min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^\top H\mathbf{x} + \mathbf{f}^\top\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U}\end{aligned}$$

with H **diagonal**

Proximal gradient method

Constrained problem:

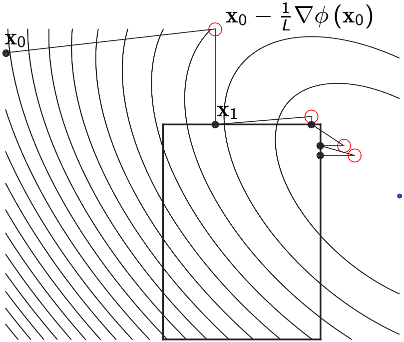
$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Special case: **clipping method**. Consider

$$\begin{aligned}\min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^\top H\mathbf{x} + \mathbf{f}^\top\mathbf{x}\\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}}\end{aligned}$$

with H **diagonal**

... has the explicit solution:

$$\mathbf{x}_i = \mathrm{prox}_X\Big(\overbrace{-H_{ii}^{-1}\mathbf{f}_i}^{2^{\mathrm{nd}}\text{-order step}}\Big)$$

Proximal gradient method

Constrained problem:

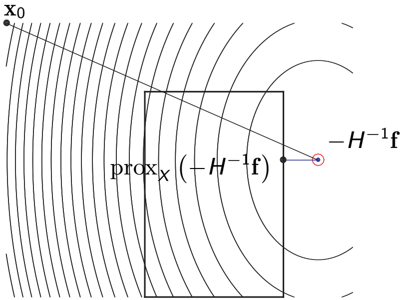
$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Special case: **clipping method**. Consider

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^\top H\mathbf{x} + \mathbf{f}^\top\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}} \end{aligned}$$

with H **diagonal**

... has the explicit solution:

$$\mathbf{x}_i = \mathrm{prox}_X\Big(\overbrace{-H_{ii}^{-1}\mathbf{f}_i}^{2^{\mathrm{nd}}\text{-order step}}\Big)$$

Proximal gradient method

Constrained problem:

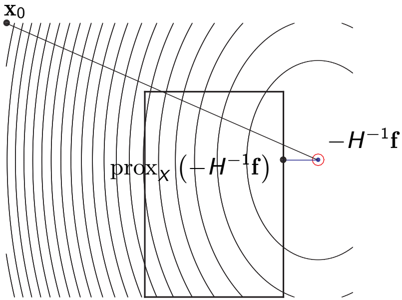
$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Special case: **clipping method**. Consider

$$\begin{aligned}\min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^\top H\mathbf{x} + \mathbf{f}^\top\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_{\mathrm{L}} \leq \mathbf{x} \leq \mathbf{x}_{\mathrm{U}}\end{aligned}$$

with H **diagonal**

... has the explicit solution:

$$\mathbf{x}_i = \mathrm{mid}\left(\mathbf{x}_{\mathrm{L},i},\ -H_{ii}^{-1}\mathbf{f}_i,\ \mathbf{x}_{\mathrm{U},i}\right)$$

where mid is the "midpoint-rule"

Proximal gradient method

Constrained problem:

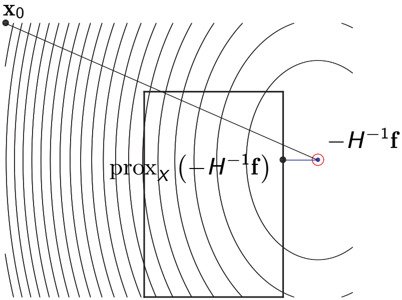
$$\min_{\mathbf{x}\in X} \quad \phi(\mathbf{x})$$

Proximal operator:

$$\mathrm{prox}_X(\mathbf{y}) = \arg\min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|^2$$

Proximal gradient iterations:

$$\mathbf{x}_{k+1} = \mathrm{prox}_X\left(\mathbf{x}_k - \frac{1}{L}\nabla\phi(\mathbf{x}_k)\right)$$



Special case: **clipping method**. Consider

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^\top H\mathbf{x} + \mathbf{f}^\top\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_\mathrm{L} \leq \mathbf{x} \leq \mathbf{x}_\mathrm{U} \end{aligned}$$

with H **diagonal**

... has the explicit solution:

$$\mathbf{x}_i = \mathrm{mid}\left(\mathbf{x}_{\mathrm{L},i},\ -H_{ii}^{-1}\mathbf{f}_i,\ \mathbf{x}_{\mathrm{U},i}\right)$$

where mid is the "midpoint-rule"

Wrap-up

- Decomposition methods used intensively in optimal control & MPC
- First-order approaches studied a lot. *Preconditioning* is often a key to efficiency (make $\nabla^2\phi$ more "round", L closer to represent the curvature in all directions)
- Second-order approaches very effective, careful about non-smooth Newton !
- QP with only simple bounds and diagonal weights (typical stage problem in MPC) has explicit solutions
- Generic framework: splitting methods