# Direct Optimal Control
# Lecture 9: Parametric Optimization


Sébastien Gros

ITK, NTNU


NTNU PhD course

# Outline

# Outline

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned}\min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0\end{aligned}$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters
- Homotopies

# Parametric optimization

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters
- Homotopies
- NLP decompositions, multi-level optimization

# Parametric optimization

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w},\mathbf{p}\right)$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$
$$\mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters
- Homotopies
- NLP decompositions, multi-level optimization
- Real-time NMPC

# Parametric optimization

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters
- Homotopies
- NLP decompositions, multi-level optimization
- Real-time NMPC

**Homotopies**: suppose

- parametric NLP is *easy* to solve for $\mathbf{p} = 0$
- parametric NLP is *what we want* to solve for $\mathbf{p} = 1$

## Parametric optimization

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

Cost & solution $\mathbf{w}$ depend on the parameters $\mathbf{p}$

**Why ?**

- Fast parametric studies
- Cheap study of effect of stochastic parameters
- Homotopies
- NLP decompositions, multi-level optimization
- Real-time NMPC

**Homotopies**: suppose

- parametric NLP is *easy* to solve for $\mathbf{p} = 0$
- parametric NLP is *what we want* to solve for $\mathbf{p} = 1$

Solve for $\mathbf{p} = 0$, keep solving while increasing $\mathbf{p}$ "slowly" always bootstrapping on previous solution

# Parametric optimization

**Parametric NLP:**

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

# Parametric optimization

**Parametric NLP:**

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

One ought to understand

$$\mathbf{w}(\mathbf{p}) : \mathbf{p} \in \mathbb{R}^p \mapsto \mathbf{w} \in \mathbb{R}^n$$

as a genuine function defined by the parametric NLP

# Parametric optimization

**Parametric NLP:**

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

One ought to understand

$$
\mathbf{w}(\mathbf{p}) : \mathbf{p} \in \mathbb{R}^p \mapsto \mathbf{w} \in \mathbb{R}^n
$$

as a genuine function defined by the parametric NLP

Solution $\mathbf{w}$ is an **implicit function** of the parameters $\mathbf{p}$. We will use IFT !

# Parametric optimization

**Parametric NLP:**

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

One ought to understand

$$\mathbf{w}(\mathbf{p}) : \mathbf{p} \in \mathbb{R}^p \mapsto \mathbf{w} \in \mathbb{R}^n$$

as a genuine function defined by the parametric NLP

Solution $\mathbf{w}$ is an **implicit function** of the parameters $\mathbf{p}$. We will use IFT !

Questions:

- Continuity & differentiability ?
- Sensitivities $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$ ?
- Predictors: using $\mathbf{w}(\mathbf{p}_0)$, what can I say about $\mathbf{w}(\mathbf{p})$ ?

$$\mathbf{w}(\mathbf{p}) \approx \mathbf{w}(\mathbf{p}_0) + \frac{\partial \mathbf{w}}{\partial \mathbf{p}}(\mathbf{p} - \mathbf{p}_0)$$

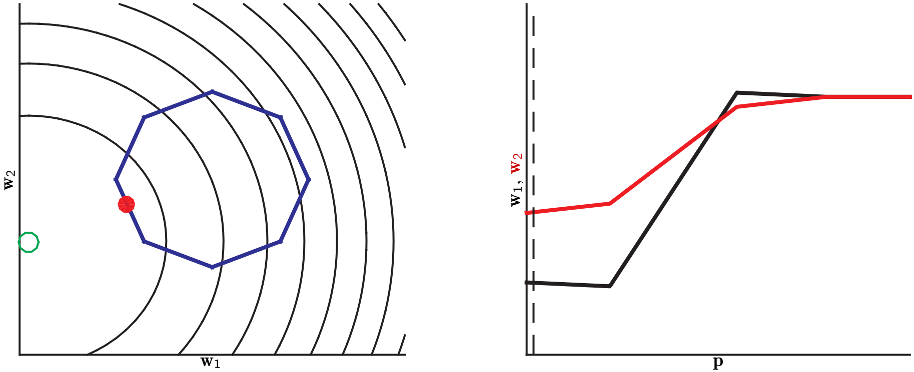
- Path-following: how to keep track of $\mathbf{w}(\mathbf{p})$ for a (continuously) changing $\mathbf{p}$ ?

# Outline

## Continuity

**Parametric NLP:**

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

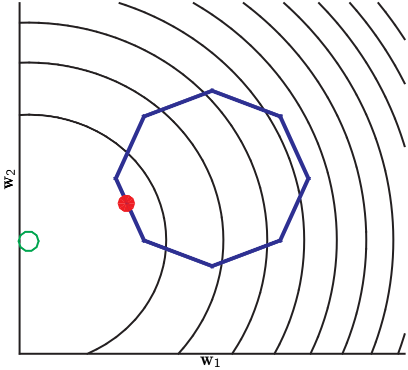


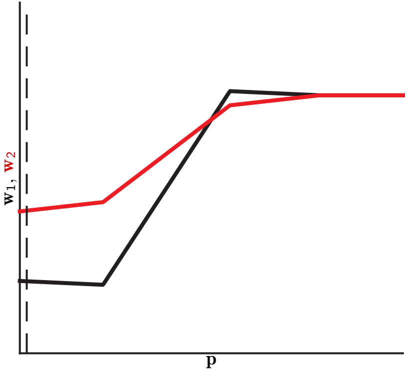
## Continuity

**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$



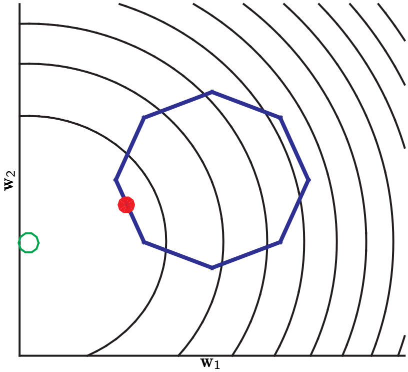
## Continuity

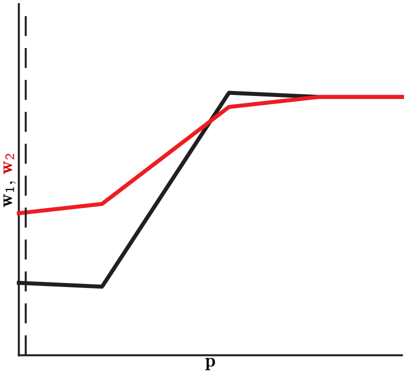
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



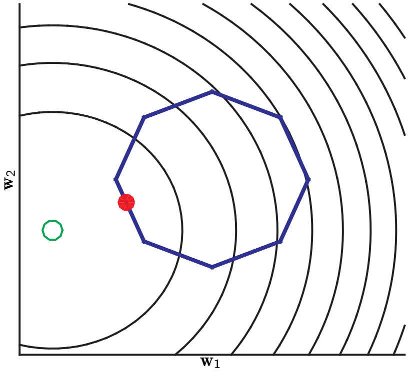
## Continuity

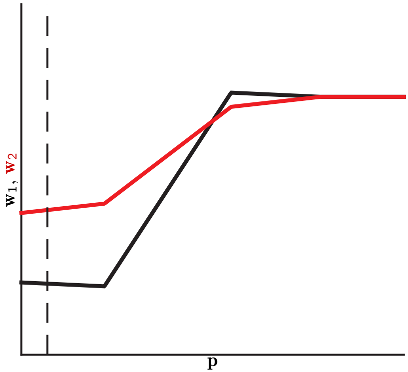
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



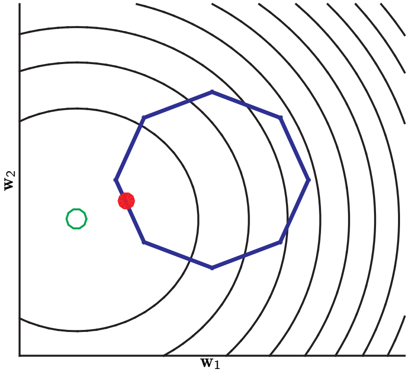
## Continuity

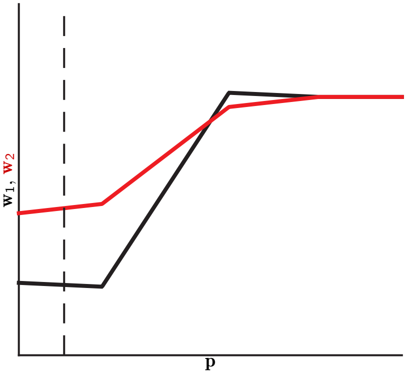
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



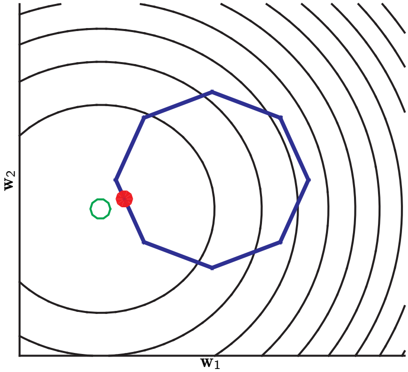
## Continuity

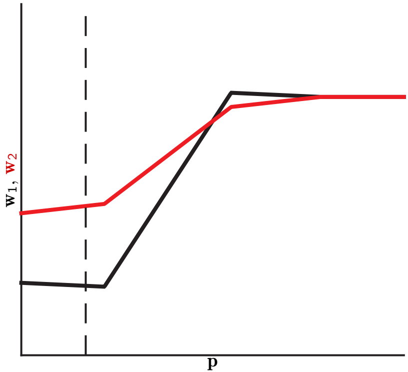
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



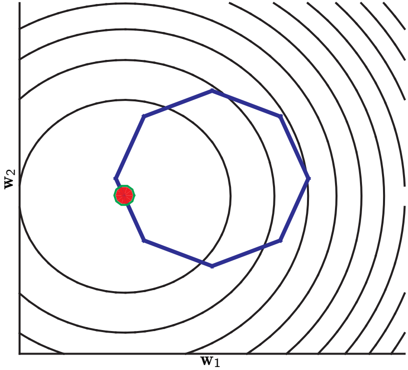
## Continuity

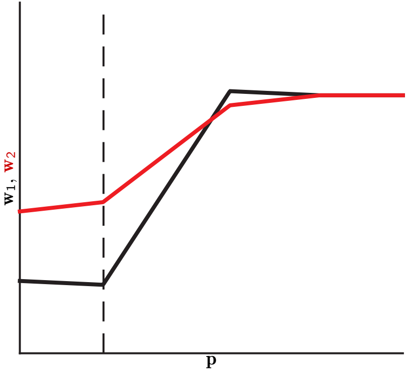
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



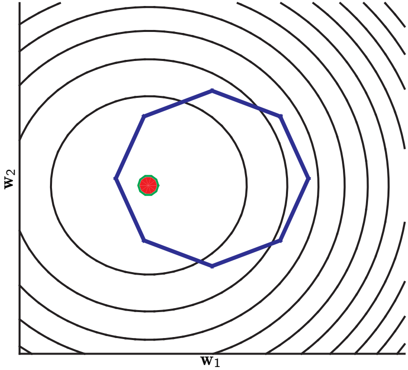
## Continuity

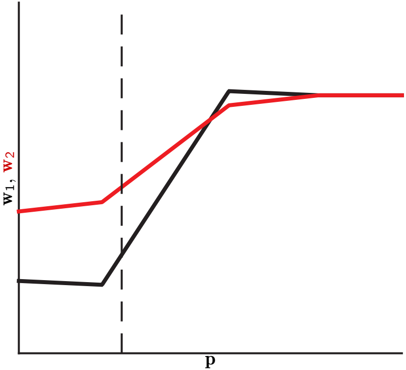
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



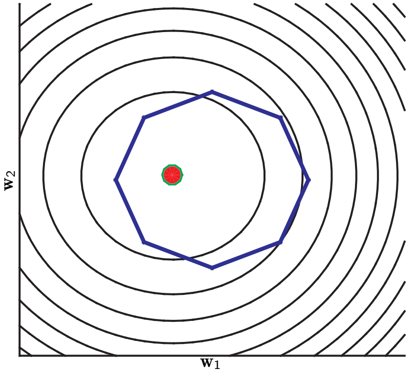
# Continuity

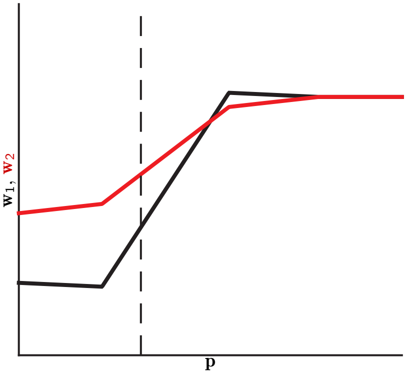
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



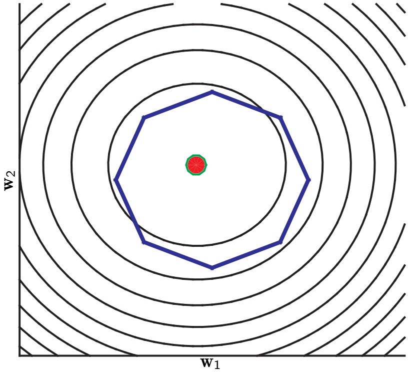
## Continuity

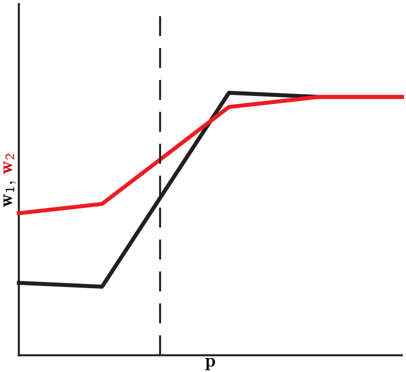
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



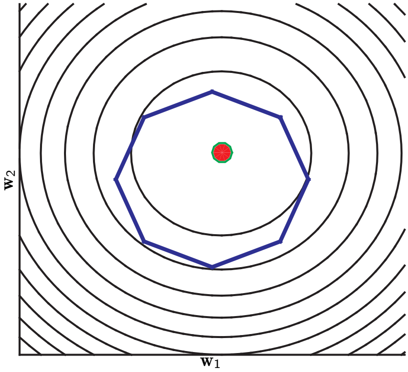
## Continuity

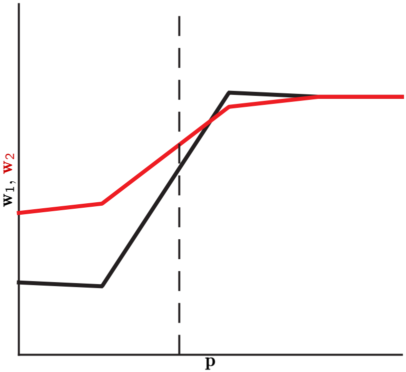
**Parametric NLP:**

$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \le 0 \end{aligned}$$

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



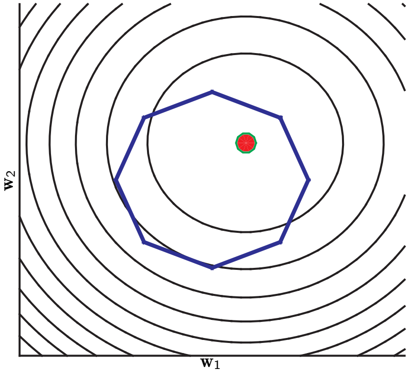
## Continuity

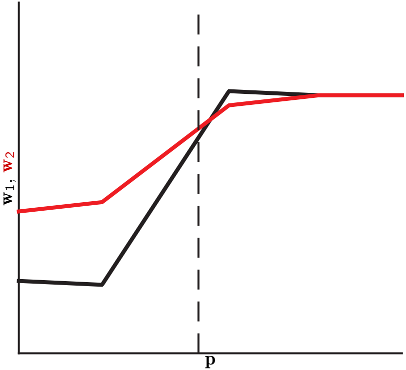
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



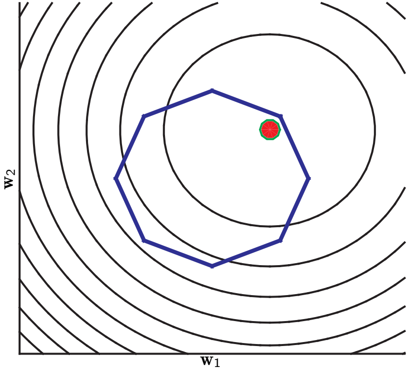
## Continuity

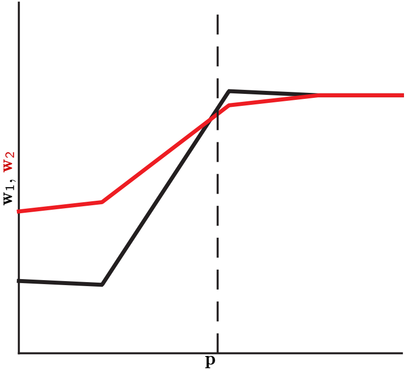
**Parametric NLP:**

$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \le 0$$

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



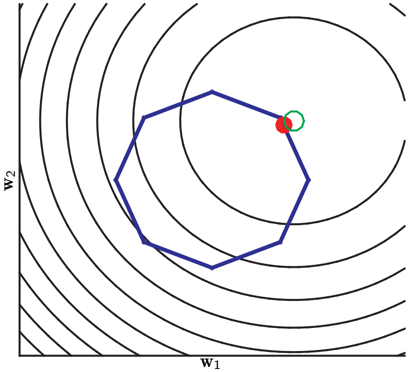
## Continuity

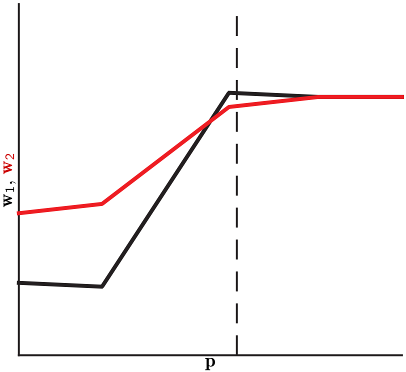
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



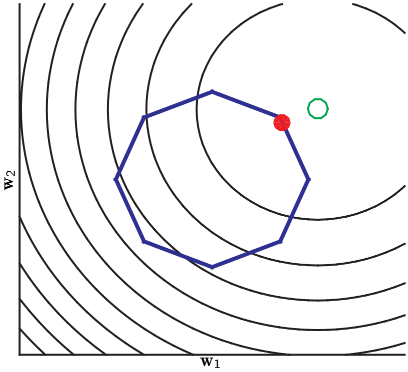
## Continuity

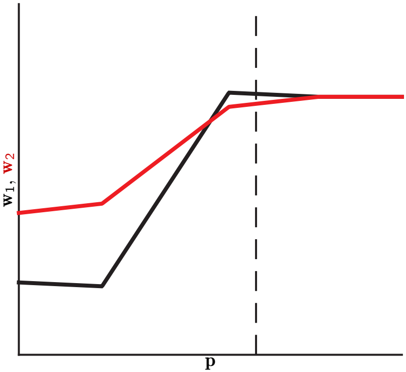
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



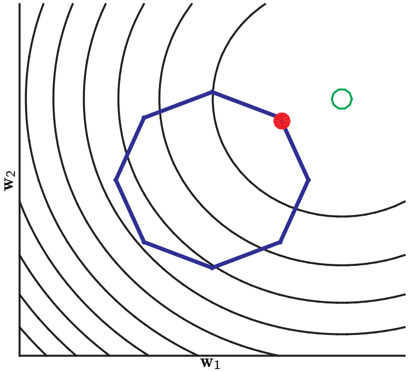
## Continuity

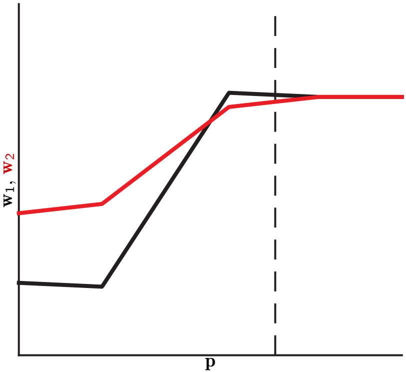
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



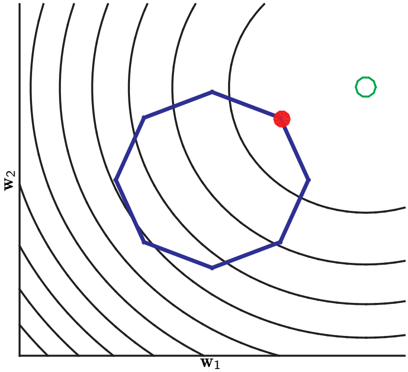
## Continuity

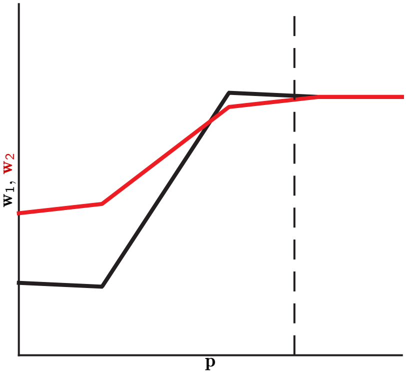
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



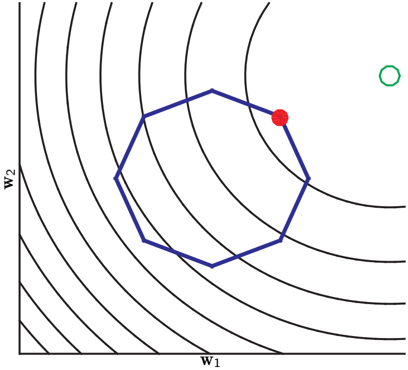
## Continuity

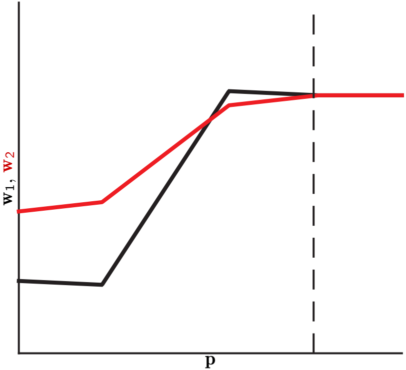
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



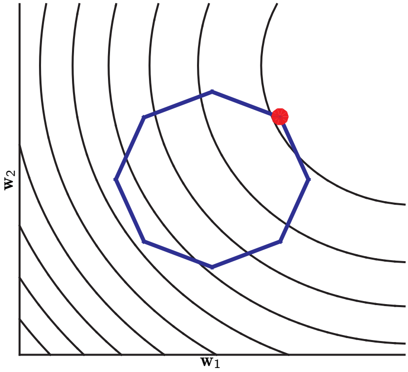
## Continuity

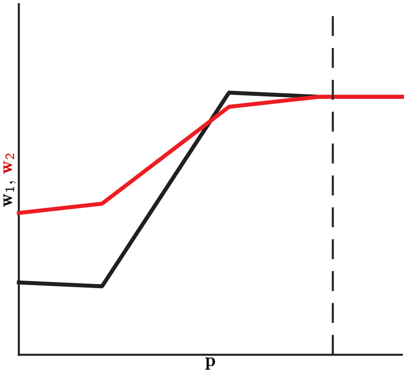
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



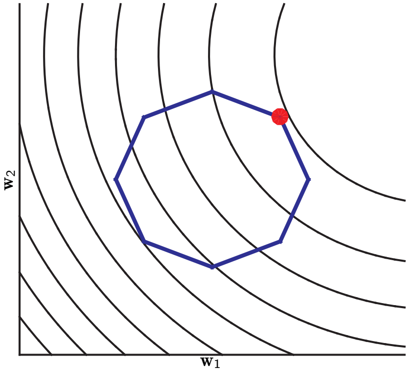
## Continuity

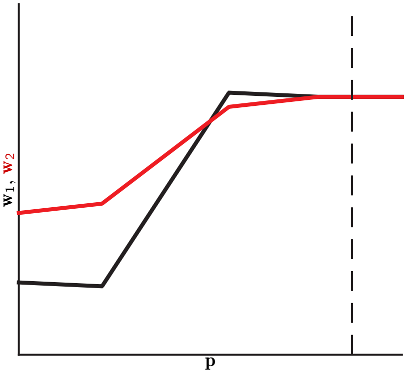
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



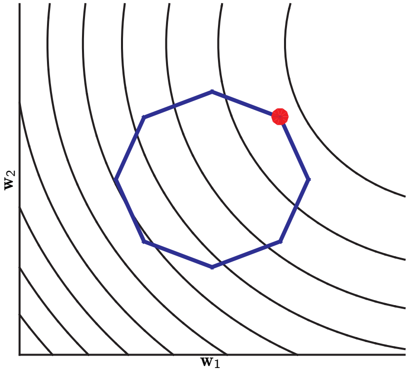
## Continuity

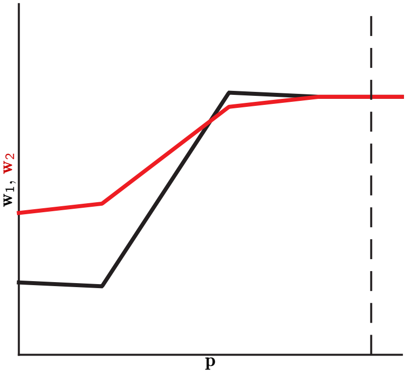
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. convex QP**: SOSC holds everywhere (positive def. Hessian)



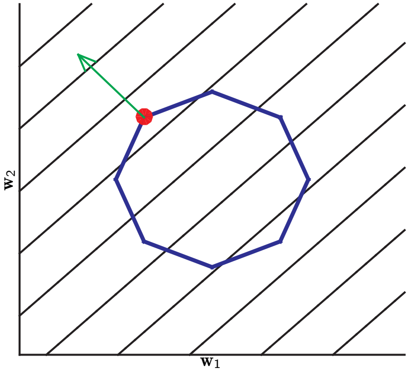
## Continuity

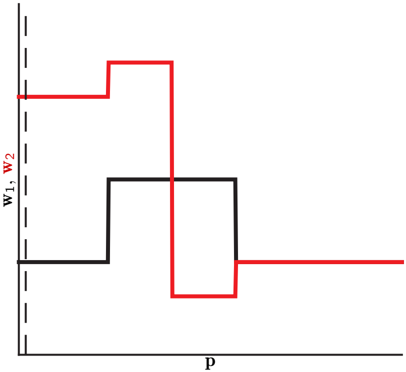
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



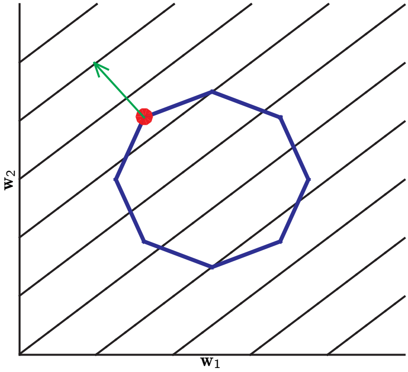
# Continuity

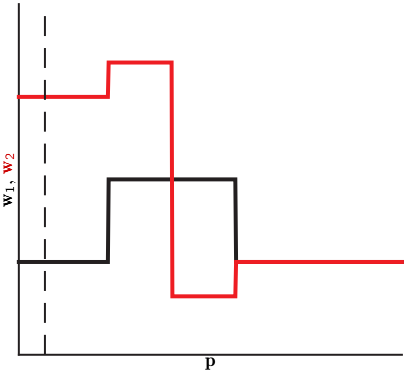
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



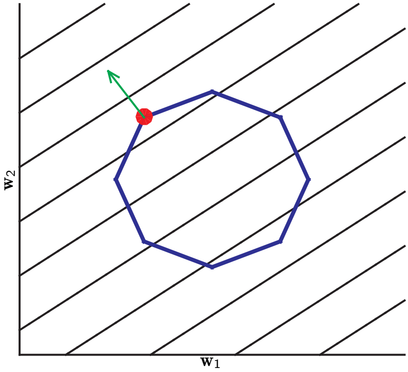
# Continuity

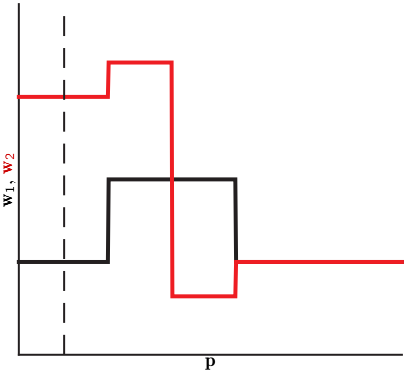
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



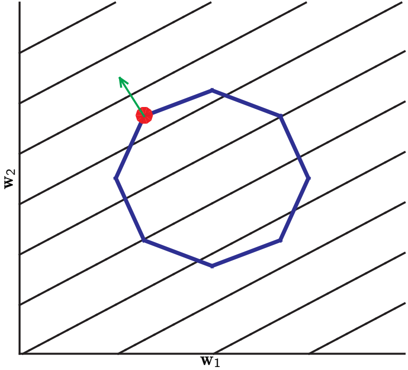
## Continuity

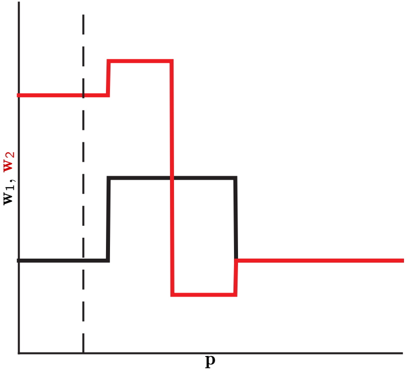
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



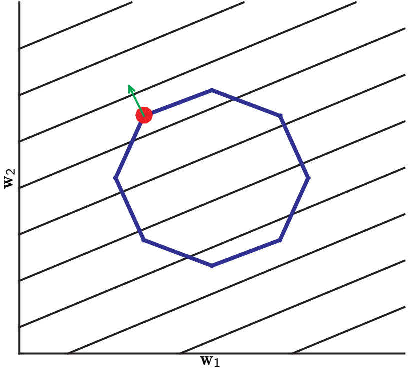
## Continuity

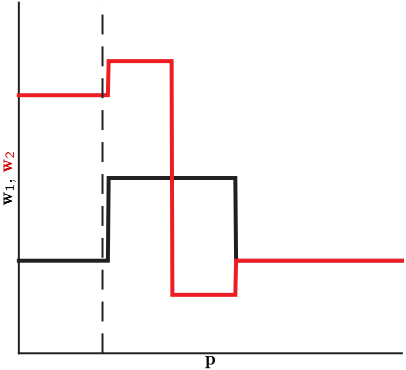
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



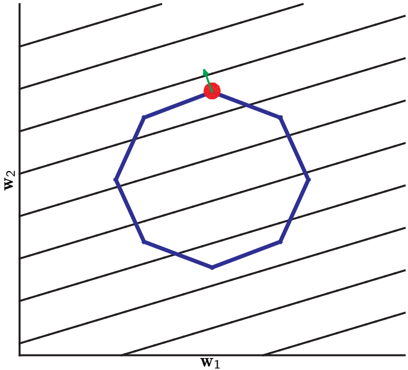
## Continuity

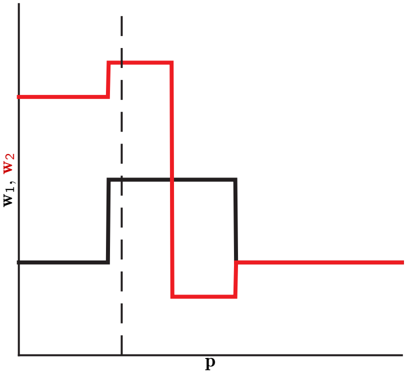
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



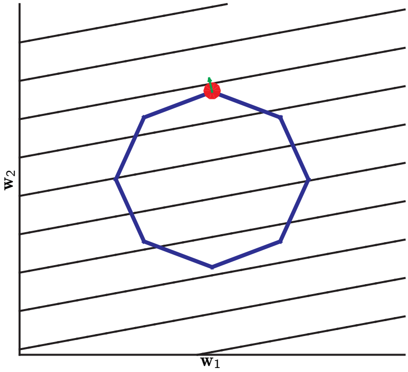
## Continuity

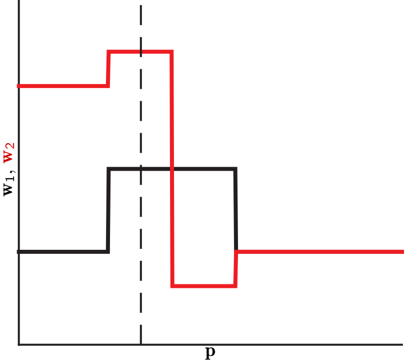
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



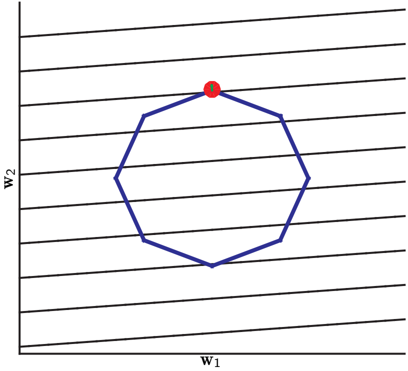
## Continuity

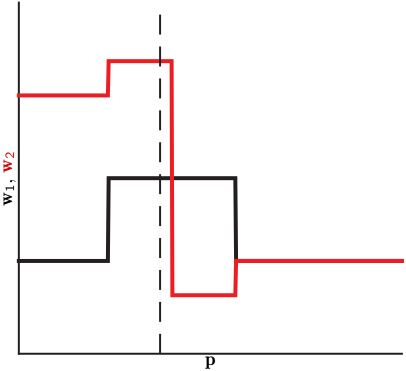
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



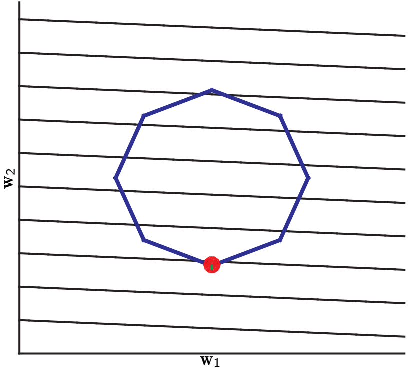
## Continuity

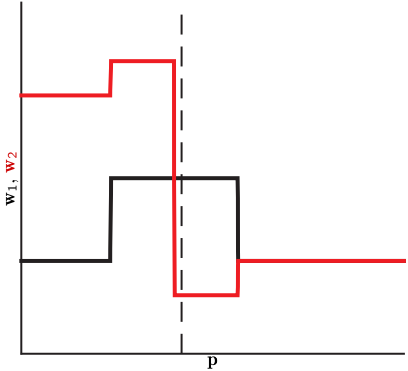
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



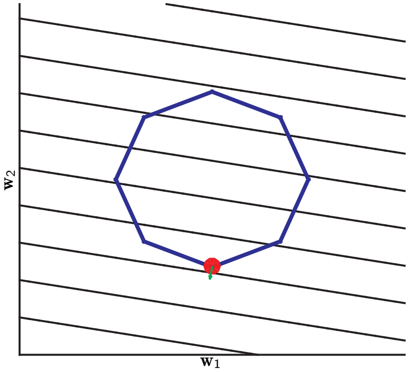
## Continuity

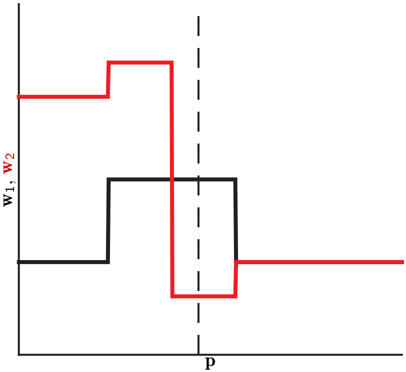
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



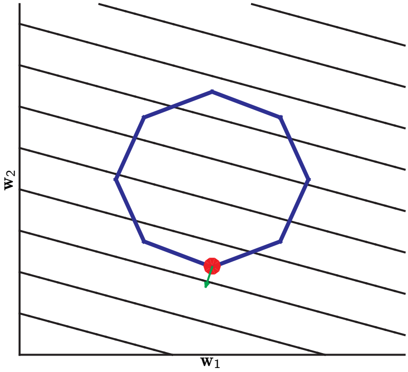
## Continuity

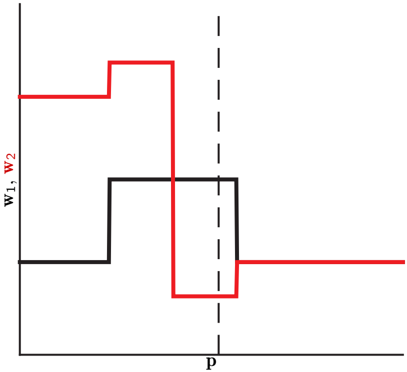
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



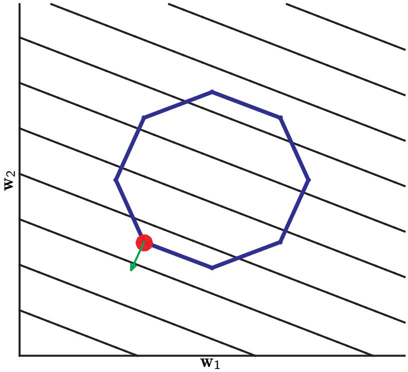
## Continuity

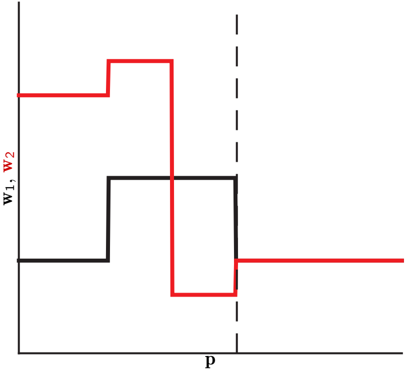
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



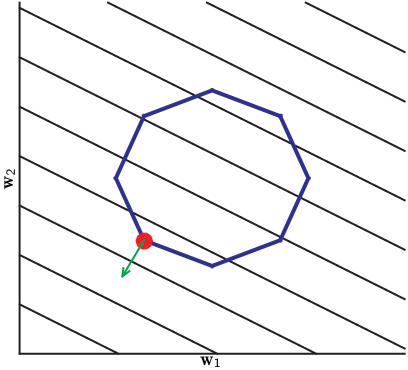
# Continuity

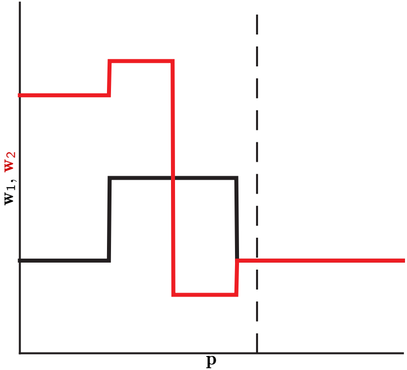
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



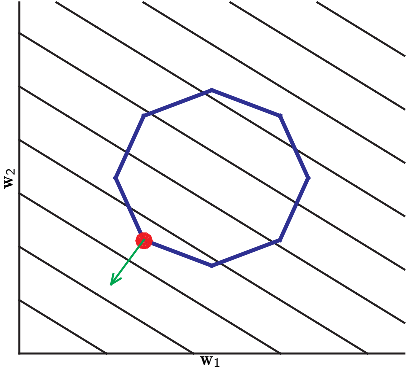
## Continuity

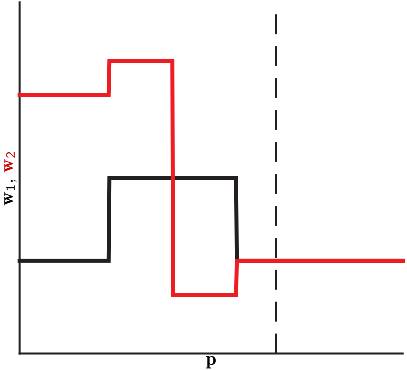
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



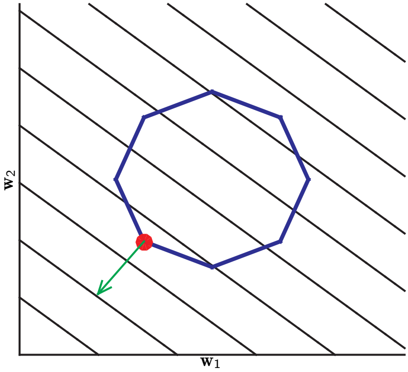
## Continuity

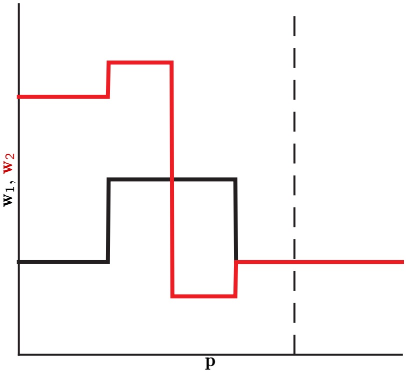
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



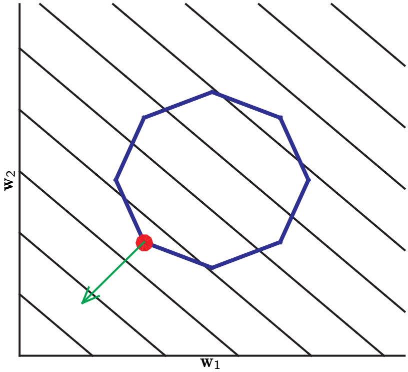
## Continuity

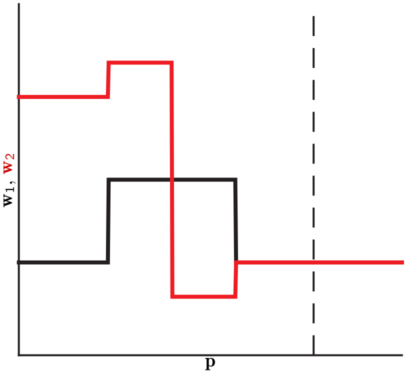
**Parametric NLP:**

$$\mathbf{w}\left(\mathbf{p}\right) = \arg\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w},\mathbf{p}\right)$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$
$$\mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0$$

**Theorem**: if $\mathbf{w}\left(\mathbf{p}\right)$ fulfils LICQ & strict SOSC, $\mathbf{w}\left(\mathbf{p}\right)$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



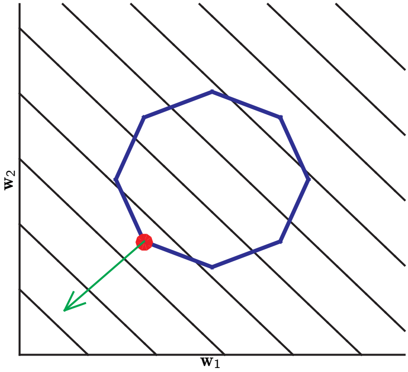
## Continuity

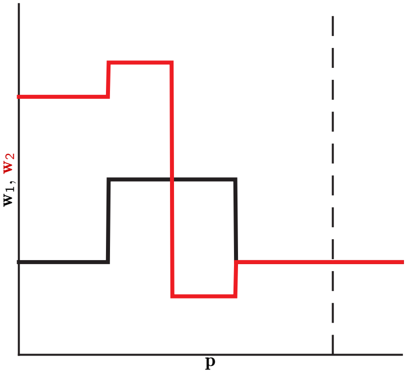
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



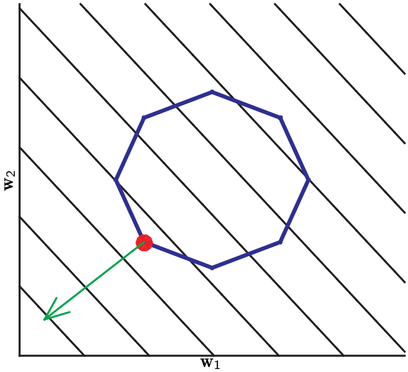
## Continuity

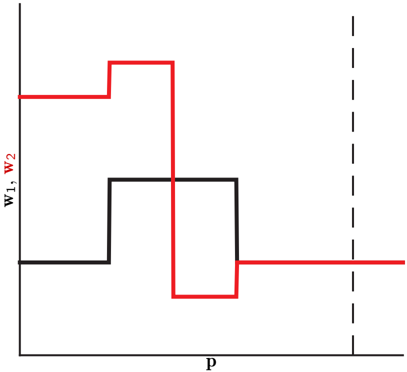
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !



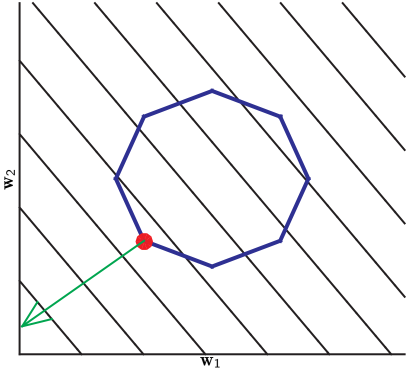
## Continuity

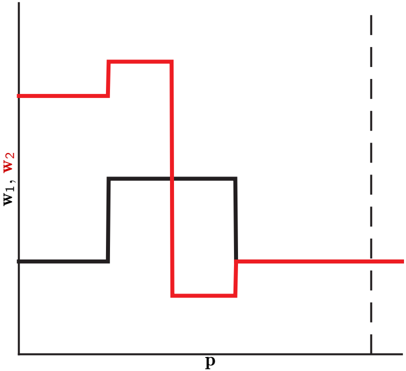
**Parametric NLP:**

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

**Theorem**: if $\mathbf{w}(\mathbf{p})$ fulfils LICQ & strict SOSC, $\mathbf{w}(\mathbf{p})$ is continuous around $\mathbf{p}$

**E.g. LP**: SOSC holds only when **2** constraints are strictly active !

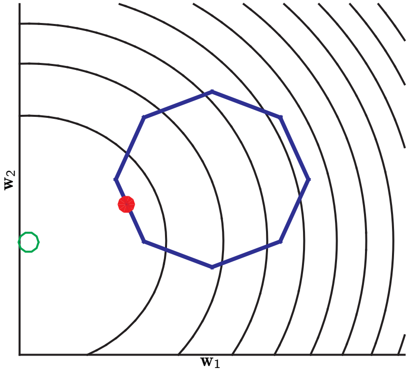


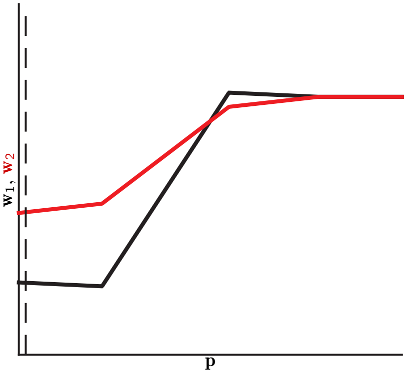



# Differentiability

**Parametric NLP:**

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

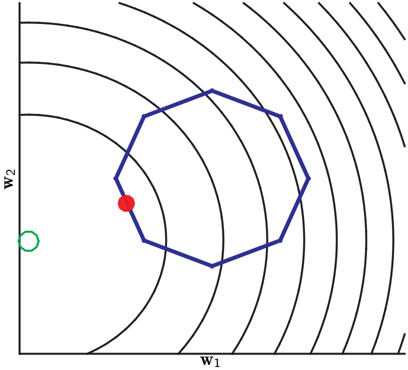


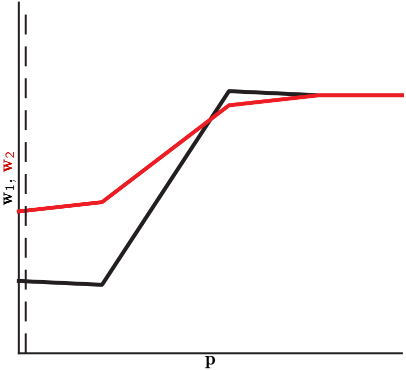



## Differentiability

**Parametric NLP:**

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



# Differentiability

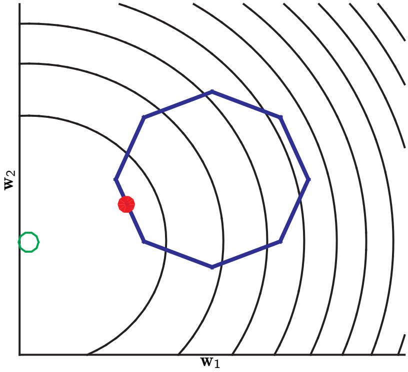
**Parametric NLP:**

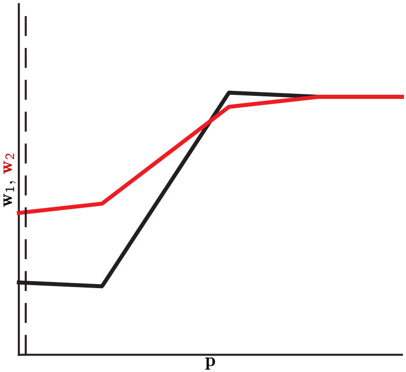
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

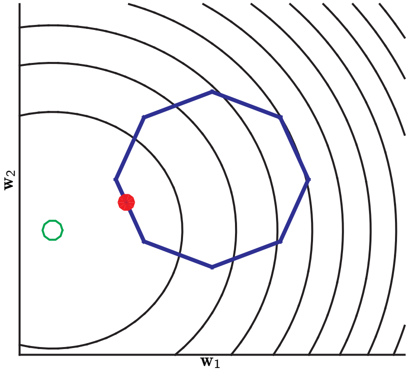
**Parametric NLP:**

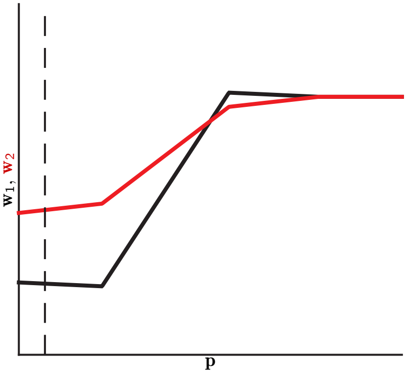
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

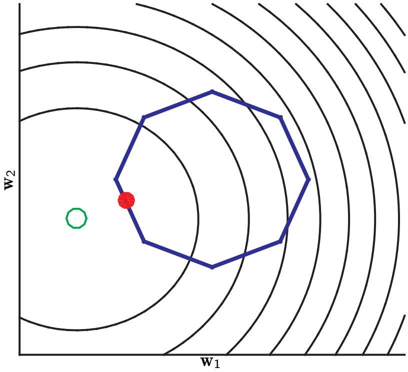
**Parametric NLP:**

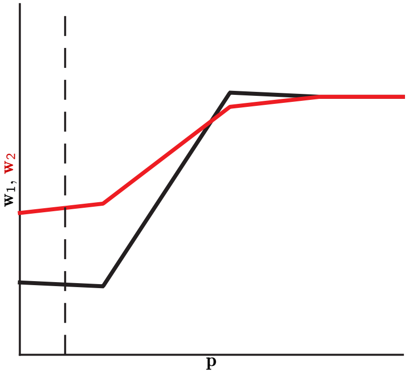
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

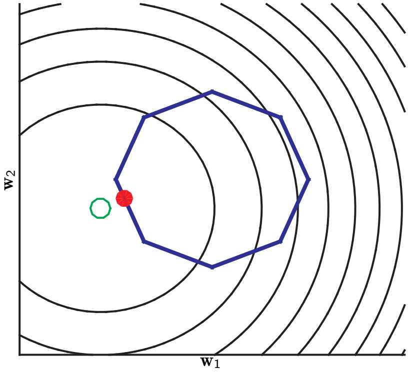
**Parametric NLP:**

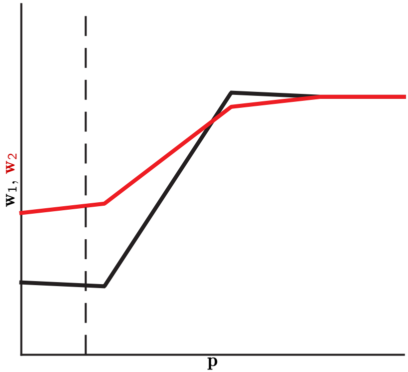
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

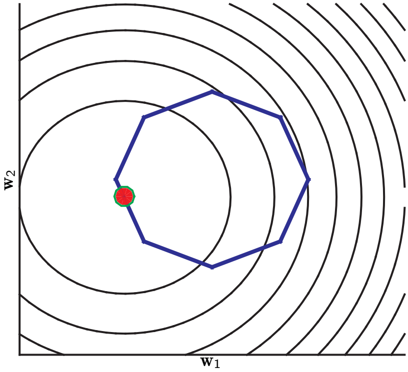
**Parametric NLP:**

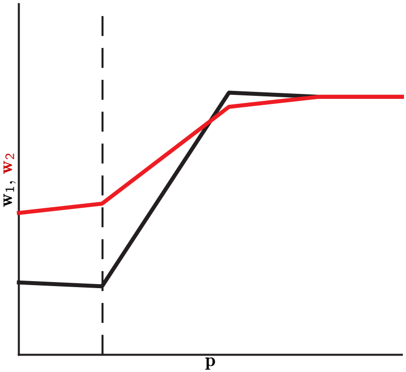
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

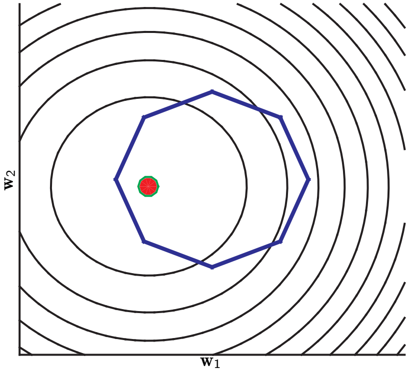
**Parametric NLP:**

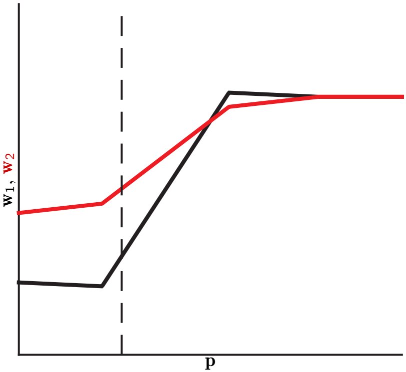
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

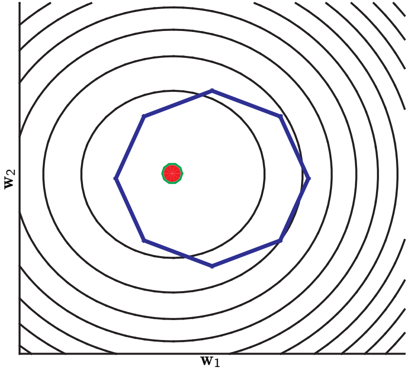
**Parametric NLP:**

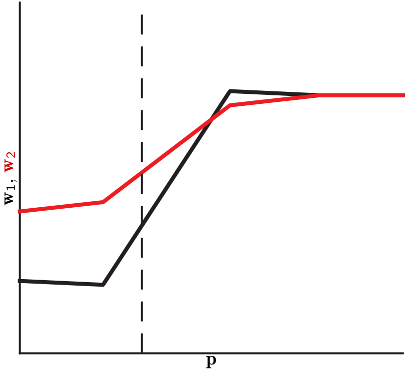
$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

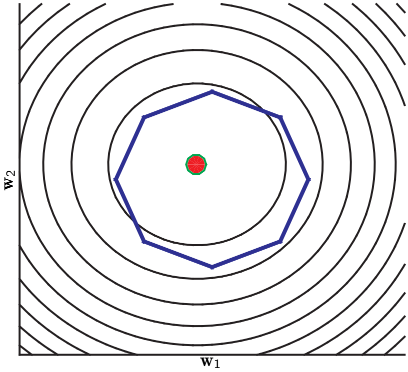
**Parametric NLP:**

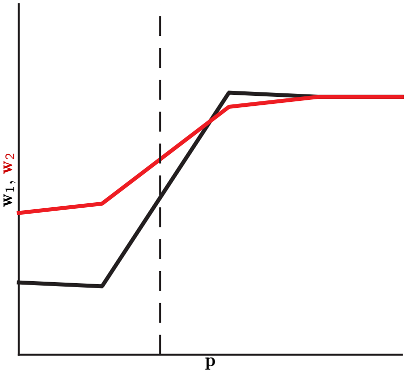
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

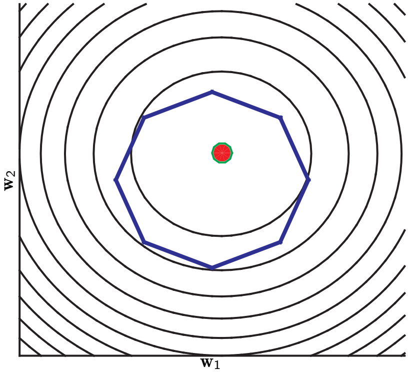
**Parametric NLP:**

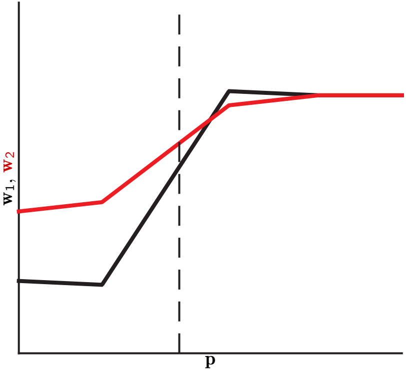
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

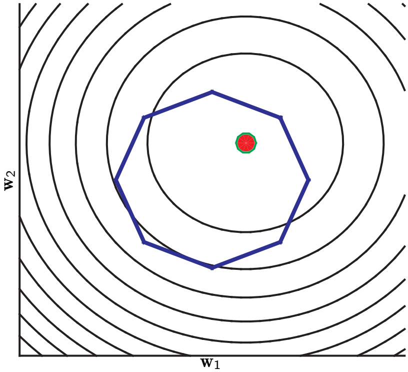
**Parametric NLP:**

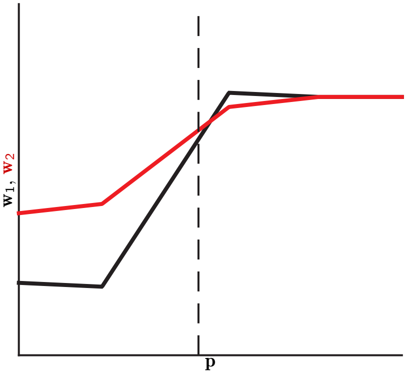
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}} \mathbf{w}(\mathbf{p})$ exists.



# Differentiability

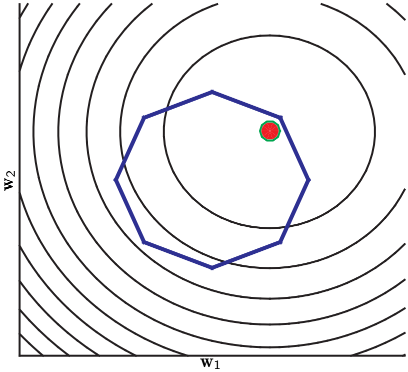
**Parametric NLP:**

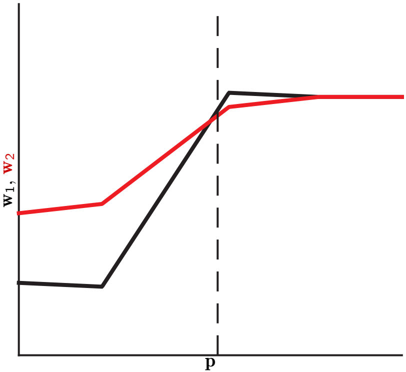
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

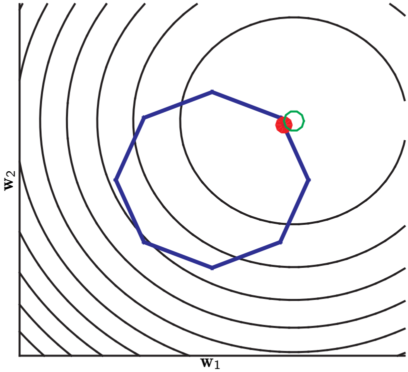
**Parametric NLP:**

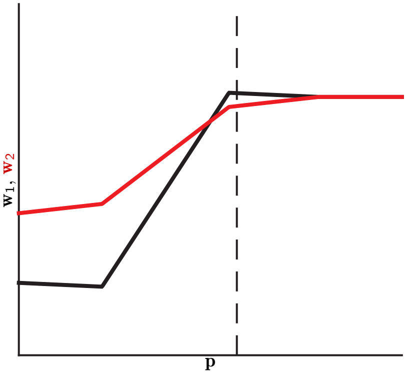
$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

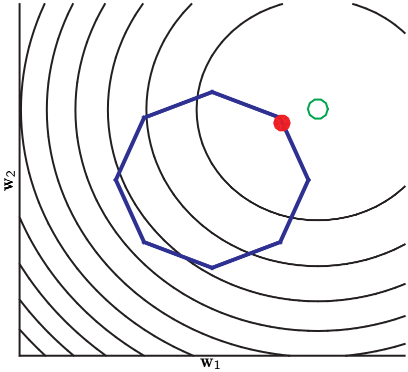
**Parametric NLP:**

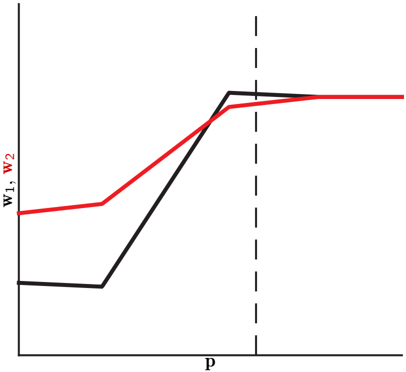
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

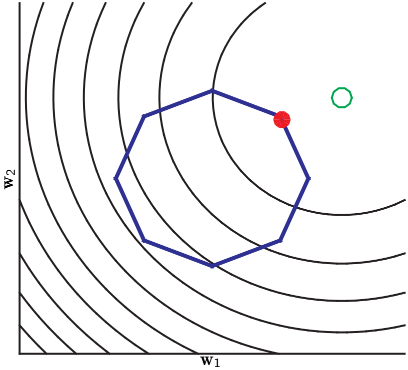
**Parametric NLP:**

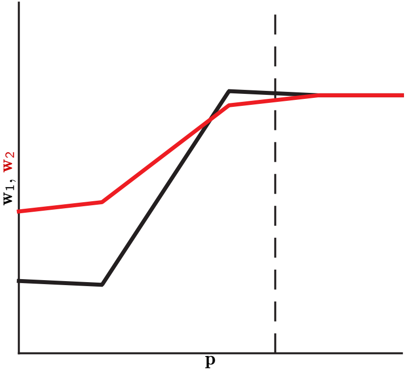
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

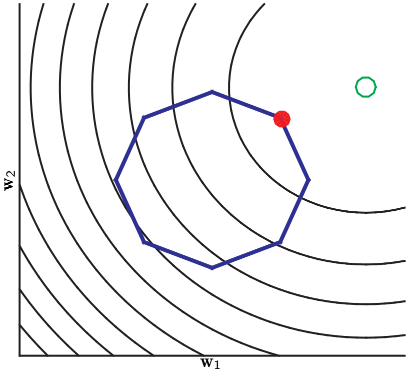
**Parametric NLP:**

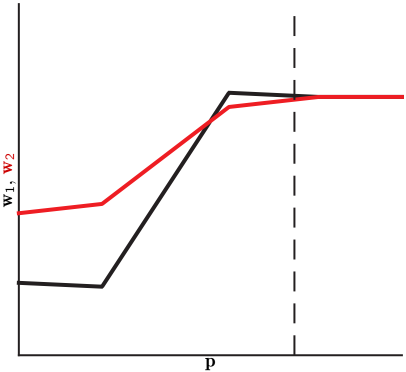
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

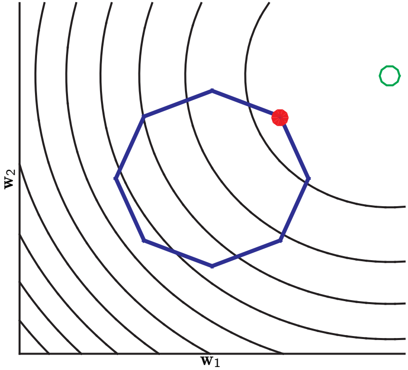
**Parametric NLP:**

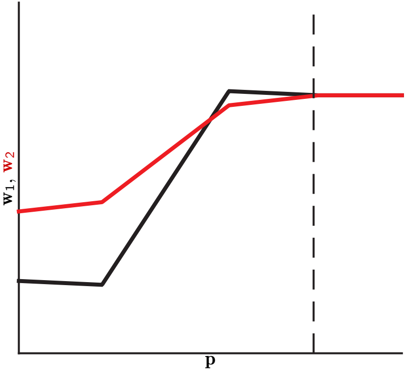
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

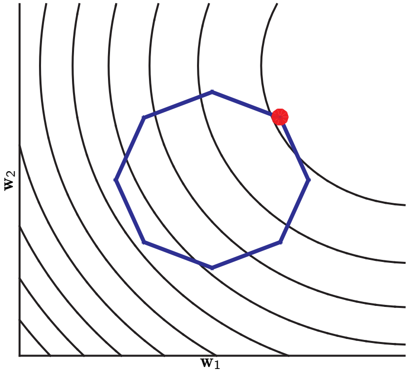
**Parametric NLP:**

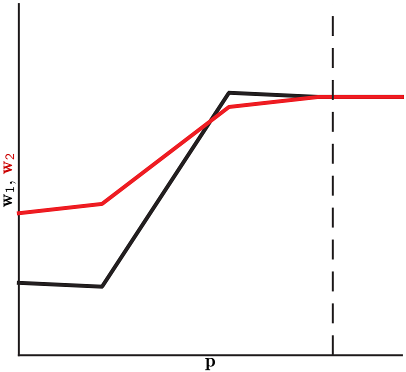
$$
\begin{aligned}
\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\
& \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\
& \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0
\end{aligned}
$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

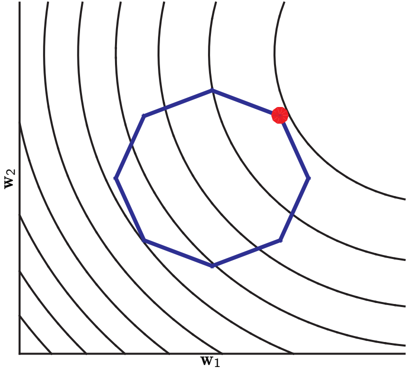
**Parametric NLP:**

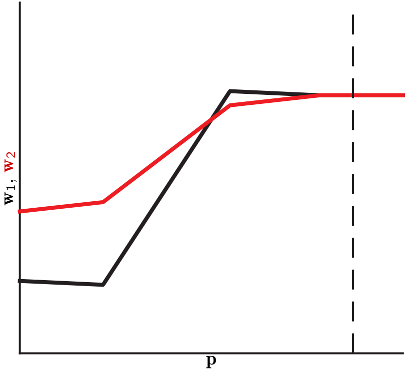
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

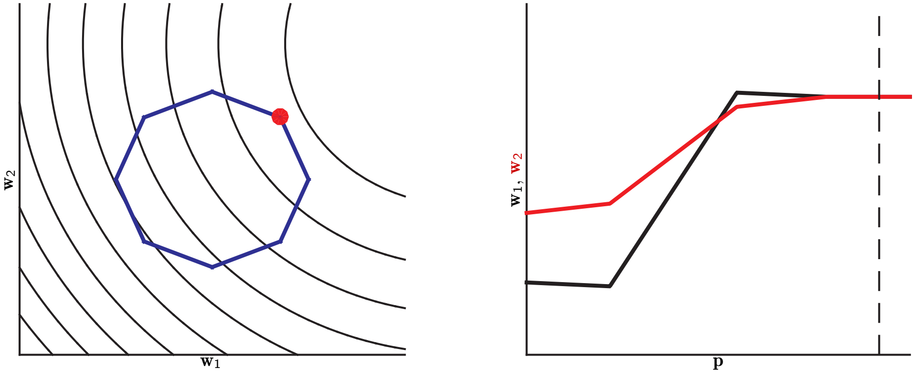
**Parametric NLP:**

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

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

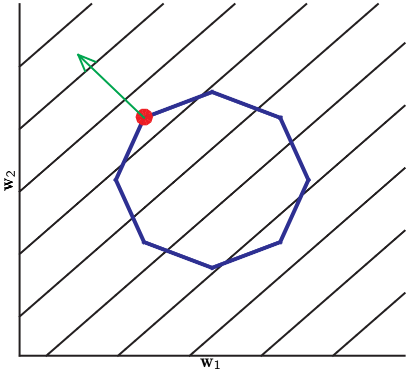
**Parametric NLP:**

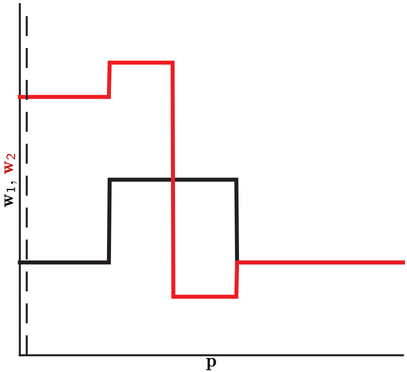
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

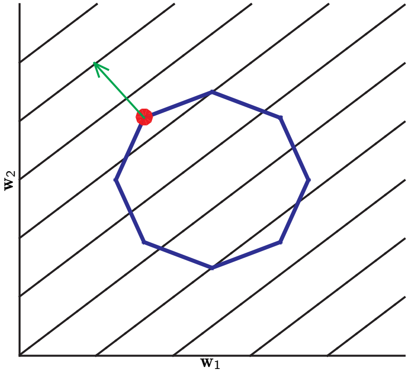
**Parametric NLP:**

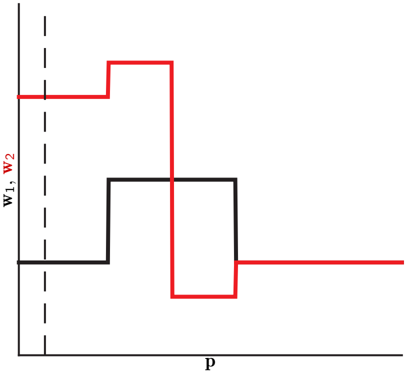
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

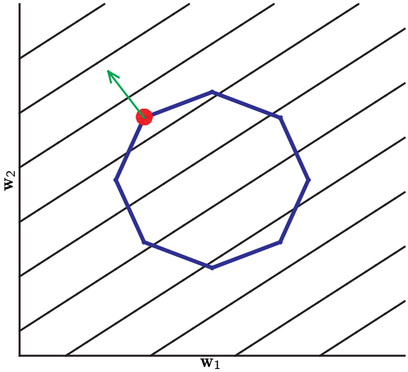
**Parametric NLP:**

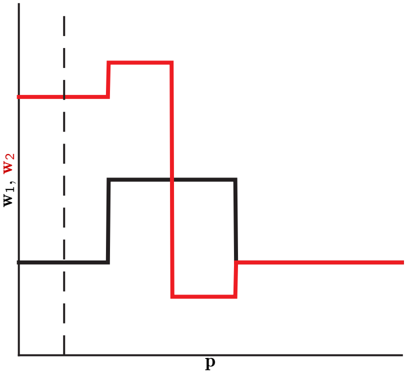
$$\begin{aligned} \mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

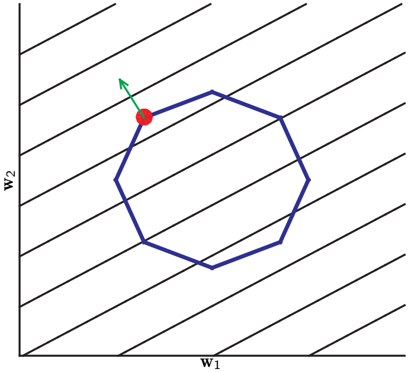
**Parametric NLP:**

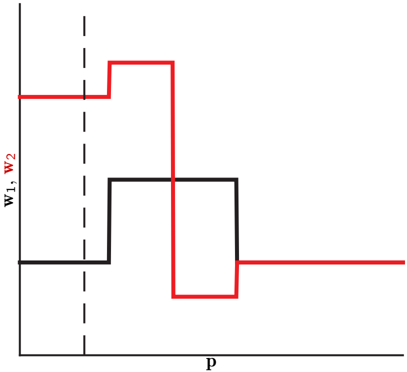
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

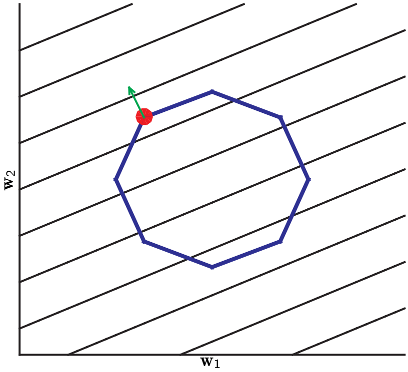
**Parametric NLP:**

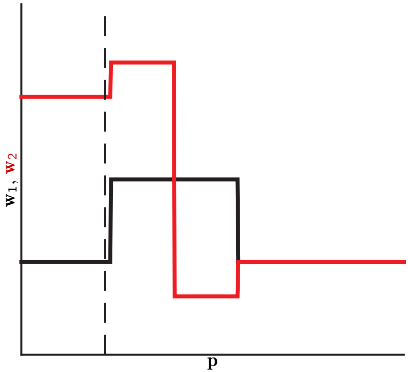
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

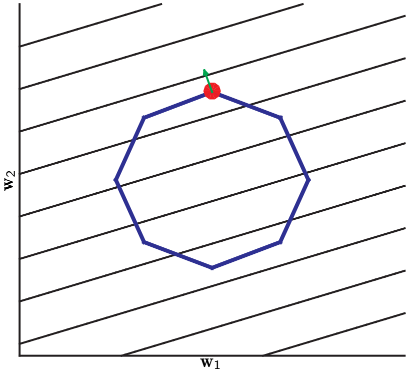
**Parametric NLP:**

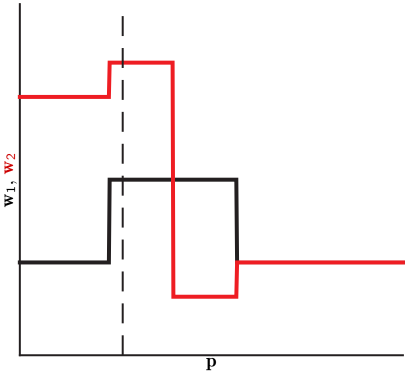
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

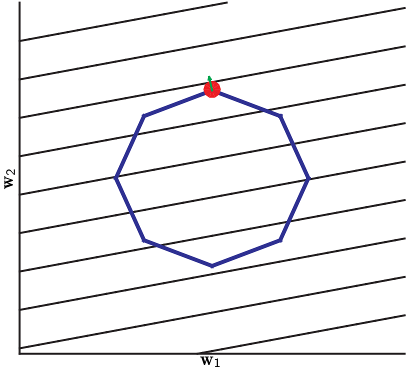
**Parametric NLP:**

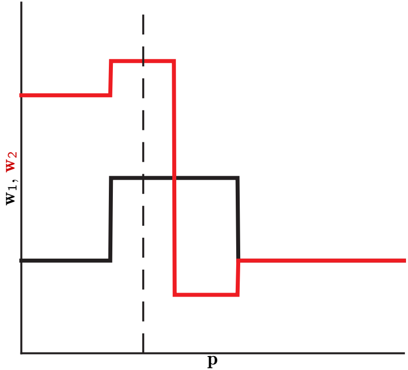
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

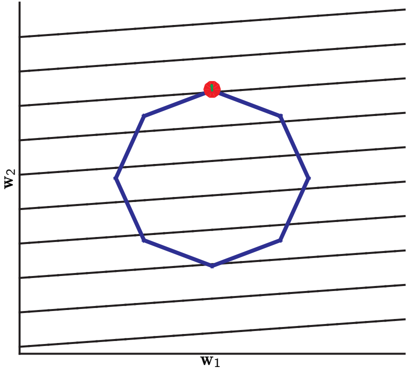
**Parametric NLP:**

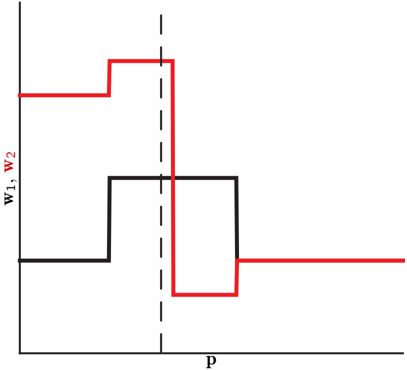
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}} \mathbf{w}(\mathbf{p})$ exists.



# Differentiability

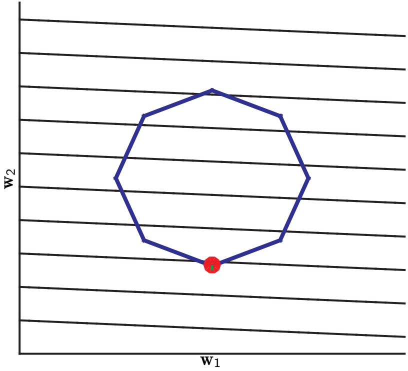
**Parametric NLP:**

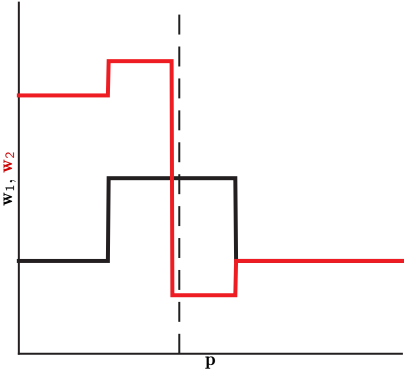
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

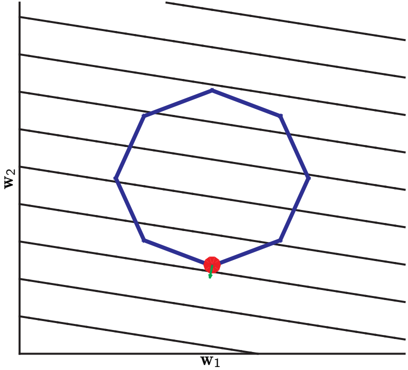
**Parametric NLP:**

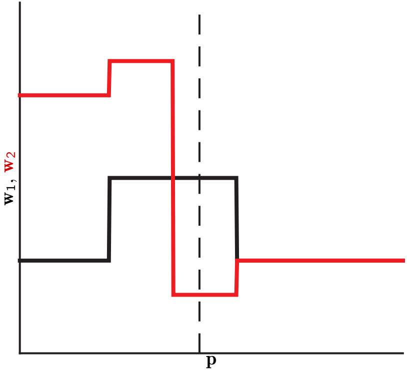
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

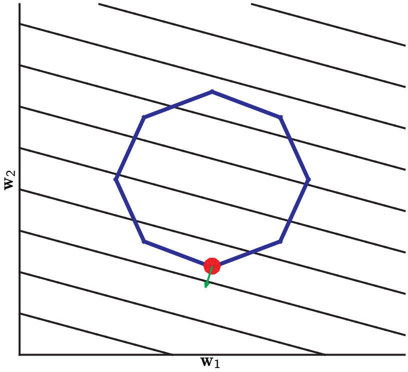
**Parametric NLP:**

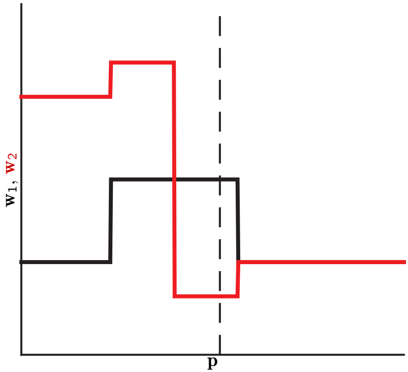
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

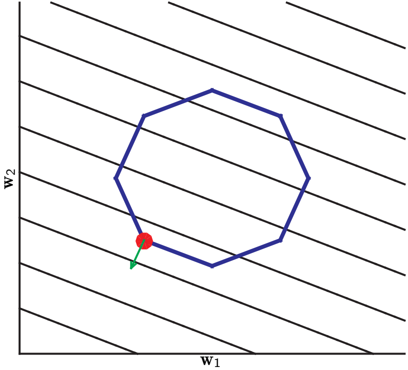
**Parametric NLP:**

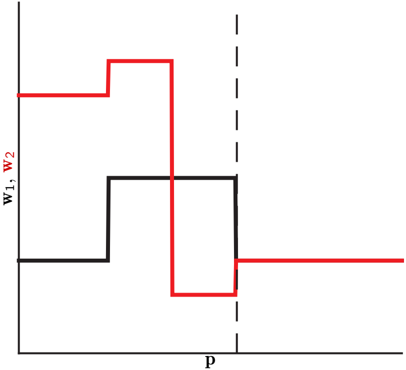
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

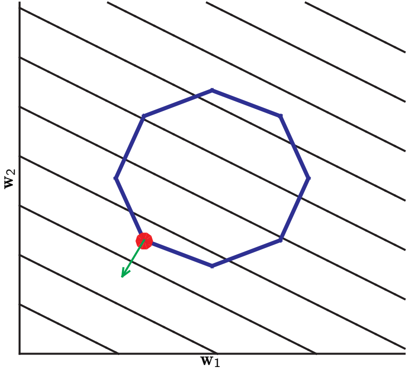
**Parametric NLP:**

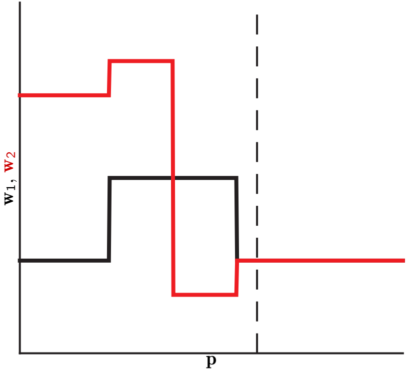
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

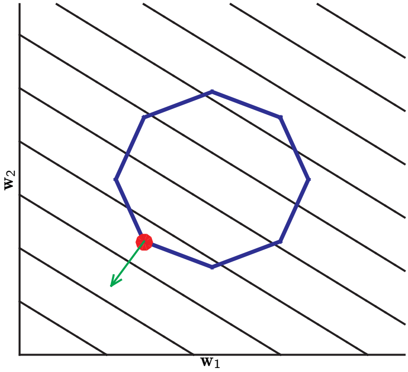
**Parametric NLP:**

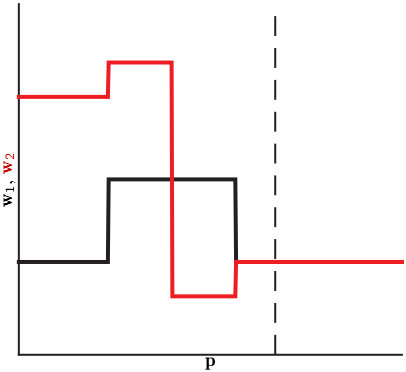
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

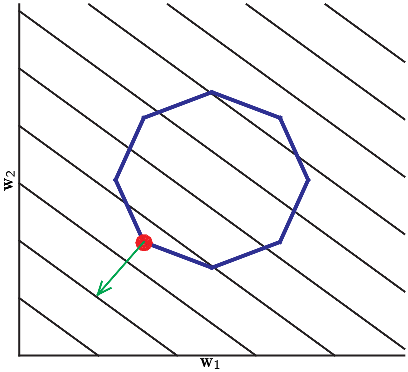
**Parametric NLP:**

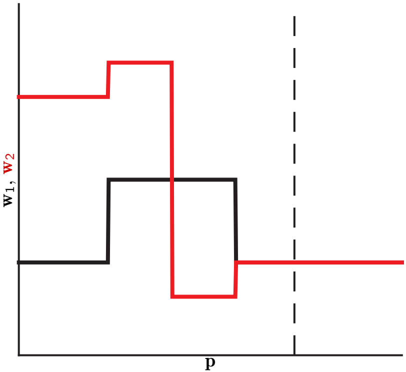
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

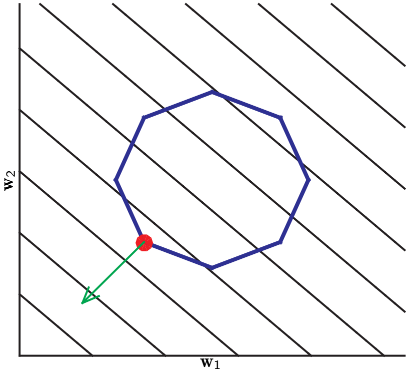
**Parametric NLP:**

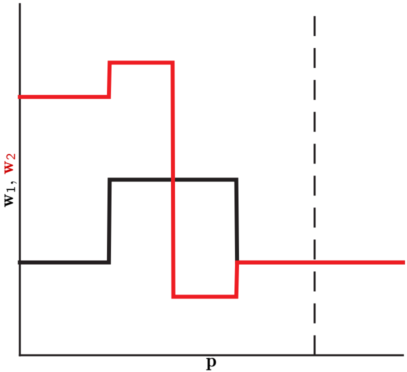
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

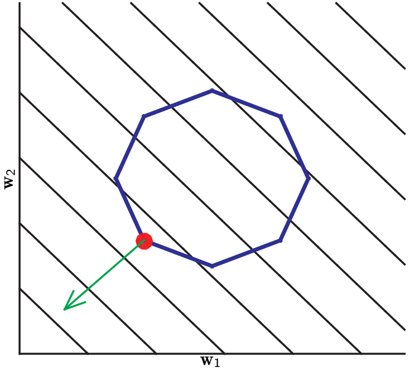
**Parametric NLP:**

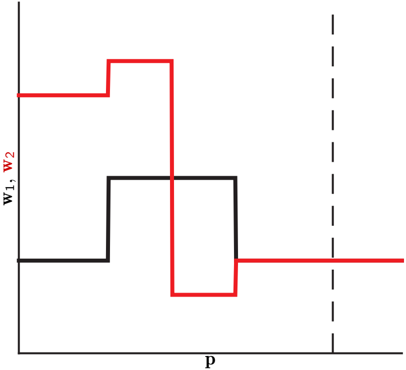
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \le 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



## Differentiability

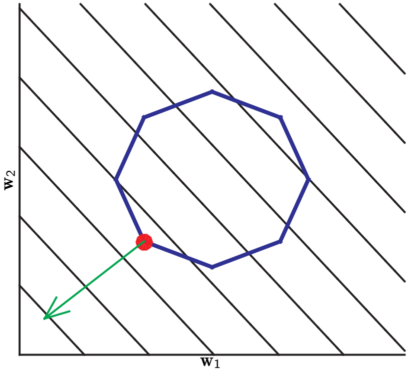
**Parametric NLP:**

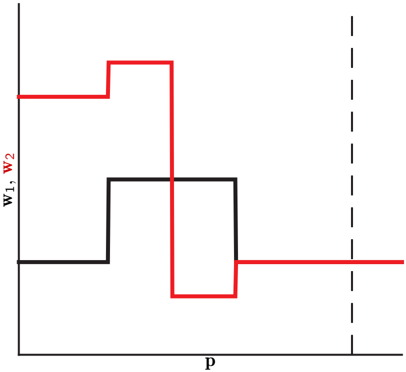
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Differentiability

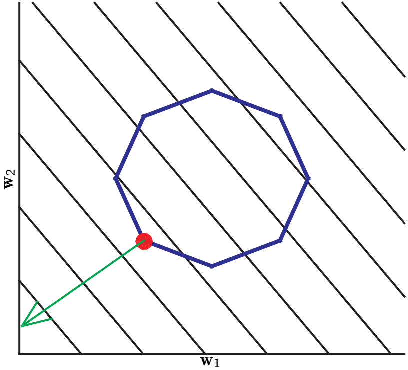
**Parametric NLP:**

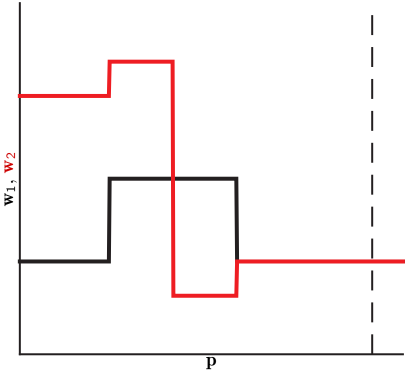
$$\mathbf{w}(\mathbf{p}) = \arg\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Theorem**: consider $\mathbf{w}(\mathbf{p})$ at a given $\mathbf{p}$, with

- LICQ & strict SOSC
- no weakly active constraint $\mathbf{h}$

then $\nabla_{\mathbf{p}}\mathbf{w}(\mathbf{p})$ exists.



# Outline

## Newton as an implicit function

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w}, \mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w}, \mathbf{p}\right) = 0 \end{aligned}$$

## Newton as an implicit function

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \end{aligned}$$

Solution $\mathbf{w}\left(\mathbf{p}\right)$, $\boldsymbol{\lambda}\left(\mathbf{p}\right)$ implicitly given by the KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) &= 0 \end{aligned}$$

## Newton as an implicit function

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0$$

# Newton as an implicit function

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

## Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0$$

We will have:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$
$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

# Newton as an implicit function

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

## Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}_0) \ \text{full rank}$$

We will have:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Newton as an implicit function

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \end{aligned}$$

### Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}_0) \text{ full rank}$$

Then for any $\mathbf{p}_0$ there is a $\mathcal{C}^1$ function $\xi(\mathbf{p})$ such that:

$$\mathbf{R}(\xi(\mathbf{p}),\mathbf{p}) = 0$$

holds in a neighbourhood of $\mathbf{p}_0$.

We will have:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Newton as an implicit function

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w},\mathbf{p}\right)$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

Solution $\mathbf{w}\left(\mathbf{p}\right), \boldsymbol{\lambda}\left(\mathbf{p}\right)$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}\right) = 0$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

### Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}\left(\mathbf{z},\mathbf{p}\right) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}\left(\mathbf{z},\mathbf{p}_0\right) \text{ full rank}$$

Then for any $\mathbf{p}_0$ there is a $\mathcal{C}^1$ function $\xi\left(\mathbf{p}\right)$ such that:

$$\mathbf{R}\left(\xi\left(\mathbf{p}\right),\mathbf{p}\right) = 0$$

holds in a neighbourhood of $\mathbf{p}_0$.

That means we have $\mathbf{z}\left(\mathbf{p}\right) = \xi\left(\mathbf{p}\right)$ around $\mathbf{p}_0$.

We will have:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \end{bmatrix}$$

## Newton as an implicit function

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Implicit function theorem

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}_0) \ \text{full rank}$$

Then for any $\mathbf{p}_0$ there is a $\mathcal{C}^1$ function $\xi(\mathbf{p})$ such that:

$$\mathbf{R}(\xi(\mathbf{p}),\mathbf{p}) = 0$$

holds in a neighbourhood of $\mathbf{p}_0$.

That means we have $\mathbf{z}(\mathbf{p}) = \xi(\mathbf{p})$ around $\mathbf{p}_0$.

We will have:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

### In other words...

... our $\mathbf{z}(\mathbf{p})$ is **locally well defined and differentiable** if $\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p})$ exists and is full rank

# Implicit functions

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Let's check $\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p})$ for the KKT conditions:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix} \quad \text{with} \quad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

# Implicit functions

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$$

$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Let's check $\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p})$ for the KKT conditions:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix} \qquad \text{with} \qquad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

We get:

$$\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla^2_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) & \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p}) \\ \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})^\top & 0 \end{bmatrix}$$

## Implicit functions

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Let's check $\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p})$ for the KKT conditions:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix} \qquad \text{with} \qquad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

We get:

$$\nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla^2_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) & \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p}) \\ \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})^\top & 0 \end{bmatrix}$$

This is the KKT matrix providing the Newton step, remember:

$$\begin{bmatrix} \nabla^2_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) & \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p}) \\ \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} \nabla\Phi(\mathbf{w}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Implicit functions

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w}\right)$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

Solution $\mathbf{w}\left(\mathbf{p}\right), \boldsymbol{\lambda}\left(\mathbf{p}\right)$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) = 0$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

Let's check $\nabla_{\mathbf{z}}\mathbf{R}\left(\mathbf{z},\mathbf{p}\right)$ for the KKT conditions:

$$\mathbf{R}\left(\mathbf{z},\mathbf{p}\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \end{bmatrix} \qquad \text{with} \qquad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

We get:

$$\nabla_{\mathbf{z}}\mathbf{R}\left(\mathbf{z},\mathbf{p}\right) = \begin{bmatrix} \nabla^2_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) & \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right)^\top & 0 \end{bmatrix}$$

This is the KKT matrix providing the Newton step, remember:

$$\begin{bmatrix} \nabla^2_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) & \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right)^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} \nabla\Phi\left(\mathbf{w}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \end{bmatrix}$$

**Theorem**

The parametric solution $\mathbf{z}\left(\mathbf{p}\right)$ is well defined and $\mathcal{C}^1$ in a neighbourhood of $\mathbf{p}$ if the KKT matrix is full rank at $\mathbf{p}$. Guaranteed if solution is LICQ & SOSC !!

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Differentiating implicit functions

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}) \text{ full rank}$$

then $\mathbf{z}(\mathbf{p})$ is well defined and $\mathcal{C}^1$.

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \mathbf{p}, \boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

### Differentiating implicit functions

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z}, \mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}} \mathbf{R}(\mathbf{z}, \mathbf{p}) \ \text{full rank}$$

then $\mathbf{z}(\mathbf{p})$ is well defined and $\mathcal{C}^1$.

Sensitivity $\frac{\partial}{\partial \mathbf{p}} \mathbf{z}(\mathbf{p})$ is given by

$$\frac{\mathrm{d}\mathbf{R}(\mathbf{z}, \mathbf{p})}{\mathrm{d}\mathbf{p}} = \frac{\partial \mathbf{R}(\mathbf{z}, \mathbf{p})}{\partial \mathbf{z}} \frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{R}(\mathbf{z}, \mathbf{p})}{\partial \mathbf{p}} = 0$$

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \mathbf{p}) \\ \mathbf{g}(\mathbf{w}, \mathbf{p}) \end{bmatrix}$$

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Differentiating implicit functions

Let $\mathbf{z}$ be implicitly given by the $\mathcal{C}^1$ function:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = 0, \quad \text{with} \quad \nabla_{\mathbf{z}}\mathbf{R}(\mathbf{z},\mathbf{p}) \ \text{full rank}$$

then $\mathbf{z}(\mathbf{p})$ is well defined and $\mathcal{C}^1$.

Sensitivity $\frac{\partial}{\partial \mathbf{p}}\mathbf{z}(\mathbf{p})$ is given by

$$\frac{\mathrm{d}\mathbf{R}(\mathbf{z},\mathbf{p})}{\mathrm{d}\mathbf{p}} = \frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}} = 0$$

i.e.

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}}$$

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p})$, $\boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$

$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$

$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Differentiating the optimal solution

Sensitivity given by:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}}$$

With

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$

$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \mathbf{p}, \boldsymbol{\lambda}) = 0$$

$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

### Differentiating the optimal solution

Sensitivity given by:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}(\mathbf{z}, \mathbf{p})}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}(\mathbf{z}, \mathbf{p})}{\partial \mathbf{p}}$$

With

$$\frac{\partial \mathbf{R}(\mathbf{z}, \mathbf{p})}{\partial \mathbf{z}} = \begin{bmatrix} H(\mathbf{w}, \boldsymbol{\lambda}, \mathbf{p}) & \nabla_{\mathbf{w}} \mathbf{g}(\mathbf{w}, \mathbf{p}) \\ \nabla_{\mathbf{w}} \mathbf{g}(\mathbf{w}, \mathbf{p})^\top & 0 \end{bmatrix}$$

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \mathbf{p}) \\ \mathbf{g}(\mathbf{w}, \mathbf{p}) \end{bmatrix}$$

## Computing the Sensitivities

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w},\mathbf{p}\right)$$

$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

Solution $\mathbf{w}\left(\mathbf{p}\right)$, $\boldsymbol{\lambda}\left(\mathbf{p}\right)$ implicitly given by the KKT conditions:

$$\nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}\right) = 0$$

$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

### Differentiating the optimal solution

Sensitivity given by:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}\left(\mathbf{z},\mathbf{p}\right)}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}\left(\mathbf{z},\mathbf{p}\right)}{\partial \mathbf{p}}$$

With

$$\frac{\partial \mathbf{R}\left(\mathbf{z},\mathbf{p}\right)}{\partial \mathbf{z}} = \left[\begin{array}{cc} H\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) & \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \nabla_{\mathbf{w}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right)^\top & 0 \end{array}\right]$$

$$\frac{\partial \mathbf{R}\left(\mathbf{z},\mathbf{p}\right)}{\partial \mathbf{p}} = \left[\begin{array}{c} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) \\ \nabla_{\mathbf{p}}\mathbf{g}\left(\mathbf{w},\mathbf{p}\right)^\top \end{array}\right]$$

Where:

$$\mathbf{z} = \left[\begin{array}{c} \mathbf{w} \\ \boldsymbol{\lambda} \end{array}\right]$$

$$\mathbf{R} = \left[\begin{array}{c} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \end{array}\right]$$

## Computing the Sensitivities

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p})$ implicitly given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p}) \qquad\qquad \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda}) = 0$$

$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \qquad\qquad \mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

### Differentiating the optimal solution

Sensitivity given by:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}}$$

With

$$\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{z}} = \begin{bmatrix} H(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) & \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p}) \\ \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})^\top & 0 \end{bmatrix}$$

$$\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}} = \begin{bmatrix} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \nabla_{\mathbf{p}}\mathbf{g}(\mathbf{w},\mathbf{p})^\top \end{bmatrix}$$

Where:

$$\mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \end{bmatrix}$$

- If $\mathbf{p}$ enters linearly in $\mathbf{g}(\mathbf{w},\mathbf{p})$, then $\frac{\partial \mathbf{R}(\mathbf{z},\mathbf{p})}{\partial \mathbf{p}} = \begin{bmatrix} 0 \\ \text{Cst.} \end{bmatrix}$
- Sensitivities are **“for free”** since a factorisation of the KKT matrix is available from the Newton algorithm

## Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w}, \mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w}, \mathbf{p}\right) = 0 \end{aligned}$$

# Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$

$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

## Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$

$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

---

**Algorithm:** NLP solution with sensitivities

---

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1}$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

---

# Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w}, \mathbf{p}\right)$$
$$\mathbf{g}\left(\mathbf{w}, \mathbf{p}\right) = 0$$

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^{\top} & 0 \end{bmatrix}^{-1}$$

Newton step

$$\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^{+} \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{bmatrix}$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

## Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$

$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1}$$

Newton step

$$\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{bmatrix}$$

Update:

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow t\boldsymbol{\lambda}^+ + (1-t)\boldsymbol{\lambda}_k$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

## Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1}$$

Newton step

$$\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{bmatrix}$$

Update:

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow t\boldsymbol{\lambda}^+ + (1-t)\boldsymbol{\lambda}_k$$

Compute sensitivities at the solution:

$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \end{bmatrix}$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

# Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi\left(\mathbf{w},\mathbf{p}\right)$$
$$\mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0$$

Remarks:

- $M$ re-used in the sensitivities, computationally cheap !!

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \left[\begin{array}{cc} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{array}\right]^{-1}$$

Newton step

$$\left[\begin{array}{c} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{array}\right] = -M\left[\begin{array}{c} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{array}\right]$$

Update:

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow t\boldsymbol{\lambda}^+ + (1-t)\boldsymbol{\lambda}_k$$

Compute sensitivities at the solution:

$$\frac{\partial}{\partial\mathbf{p}}\left[\begin{array}{c} \mathbf{w} \\ \boldsymbol{\lambda} \end{array}\right] = -M\left[\begin{array}{c} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \end{array}\right]$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial\mathbf{w}}{\partial\mathbf{p}}$, $\frac{\partial\boldsymbol{\lambda}}{\partial\mathbf{p}}$

# Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$

Remarks:

- $M$ re-used in the sensitivities, computationally cheap !!
- Sensitivities are inexact if Newton is not properly converged. There are tricks to handle this problem though!!

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1}$$

Newton step

$$\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{bmatrix}$$

Update:

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow t\boldsymbol{\lambda}^+ + (1-t)\boldsymbol{\lambda}_k$$

Compute sensitivities at the solution:

$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -M \begin{bmatrix} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \end{bmatrix}$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial \mathbf{w}}{\partial \mathbf{p}}$, $\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{p}}$

## Computing the Sensitivities - Implementation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$

Remarks:

- $M$ re-used in the sensitivities, computationally cheap !!
- Sensitivities are inexact if Newton is not properly converged. There are tricks to handle this problem though!!
- Must use $\nabla_{\mathbf{w},\mathbf{p}}\mathcal{L}$ and not $\nabla_{\mathbf{w},\mathbf{p}}\Phi$ in sensitivities !

---

**Algorithm:** NLP solution with sensitivities

**Input:** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\mathbf{p}$

**while** *not converged* **do**

Compute:

$$M = \left[\begin{array}{cc} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{array}\right]^{-1}$$

Newton step

$$\left[\begin{array}{c} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{array}\right] = -M\left[\begin{array}{c} \nabla_{\mathbf{w}}\Phi \\ \mathbf{g} \end{array}\right]$$

Update:

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow t\boldsymbol{\lambda}^+ + (1-t)\boldsymbol{\lambda}_k$$

Compute sensitivities at the solution:

$$\frac{\partial}{\partial\mathbf{p}}\left[\begin{array}{c} \mathbf{w} \\ \boldsymbol{\lambda} \end{array}\right] = -M\left[\begin{array}{c} \nabla_{\mathbf{w},\mathbf{p}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \end{array}\right]$$

**return** $\mathbf{w}$, $\boldsymbol{\lambda}$, $\frac{\partial\mathbf{w}}{\partial\mathbf{p}}$, $\frac{\partial\boldsymbol{\lambda}}{\partial\mathbf{p}}$

# Outline

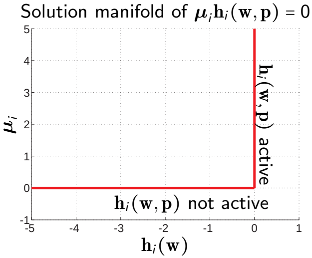
# Sensitivity with inequality constraints

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p}), \boldsymbol{\mu}(\mathbf{p})$ given by the KKT conditions:

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

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu} \geq 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) &\leq 0 \end{aligned}$$

Now we have non-smooth conditions...



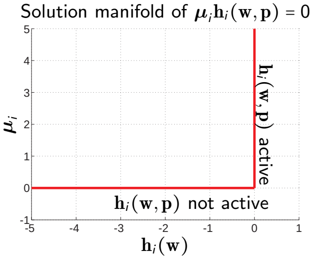
# Sensitivity with inequality constraints

**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p}), \boldsymbol{\mu}(\mathbf{p})$ given by the KKT conditions:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

$$\begin{aligned} \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w}, \mathbf{p}) &= 0 \\ \boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w}, \mathbf{p}) &= 0 \\ \boldsymbol{\mu} \geq 0, \quad \mathbf{h}(\mathbf{w}, \mathbf{p}) &\leq 0 \end{aligned}$$

Now we have non-smooth conditions...




... however, they are piecewise smooth !!

## Sensitivity with inequality constraints

**Parametric NLP:**

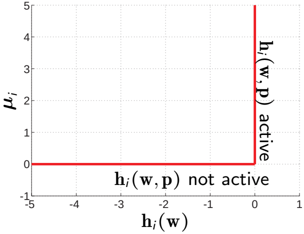
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p}), \boldsymbol{\mu}(\mathbf{p})$ given by the KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu} \geq 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) &\leq 0 \end{aligned}$$

Now we have non-smooth conditions...

Solution manifold of $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w},\mathbf{p}) = 0$




... however, they are piecewise smooth !!

Let $\mathbb{A}$ be the active set, then we have:

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0 \end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

## Sensitivity with inequality constraints

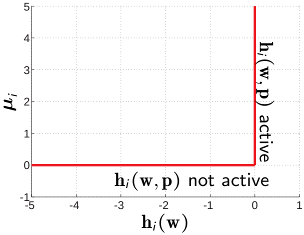
**Parametric NLP:** Solution $\mathbf{w}(\mathbf{p}), \boldsymbol{\lambda}(\mathbf{p}), \boldsymbol{\mu}(\mathbf{p})$ given by the KKT conditions:

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

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu} \geq 0, \quad \mathbf{h}(\mathbf{w},\mathbf{p}) &\leq 0 \end{aligned}$$

Now we have non-smooth conditions...

Solution manifold of $\boldsymbol{\mu}_i \mathbf{h}_i(\mathbf{w},\mathbf{p}) = 0$




... however, they are piecewise smooth !!

Let $\mathbb{A}$ be the active set, then we have:

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0 \end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

Conditions are smooth as long as all constraints are **strictly active**. Otherwise we avoid the "corner" of the complementarity slackness manifold.

# Sensitivity with inequality constraints

**Parametric NLP:**

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

Let $\mathbb{A}$ be the (strictly) active set, then we have:

$$\begin{aligned}\nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0\end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

# Sensitivity with inequality constraints

**Parametric NLP:**

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

Let $\mathbb{A}$ be the (strictly) active set, then we have:

$$\begin{aligned}\nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0\end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

Let's define:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu}_{\mathbb{A}} \end{bmatrix}$$

# Sensitivity with inequality constraints

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

Let $\mathbb{A}$ be the (strictly) active set, then we have:

$$\begin{aligned} \nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0 \end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

Let's define:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu}_{\mathbb{A}} \end{bmatrix}$$

Sensitivity given by $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$ with:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{z}} = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}^\top & 0 & 0 \end{bmatrix}$$

## Sensitivity with inequality constraints

**Parametric NLP:**

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

Let $\mathbb{A}$ be the (strictly) active set, then we have:

$$\begin{aligned} \nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0 \end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

Let's define:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu}_{\mathbb{A}} \end{bmatrix}$$

Sensitivity given by $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$ with:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{z}} = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}^\top & 0 & 0 \end{bmatrix}$$

In SQP methods, matrix $\frac{\partial \mathbf{R}}{\partial \mathbf{z}}$ inside underlying Active-Set QP solvers. I.e. we get the sensitivities for free$^\dagger$ in SQP !!

## Sensitivity with inequality constraints

**Parametric NLP:**

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

Let $\mathbb{A}$ be the (strictly) active set, then we have:

$$\begin{aligned}\nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p}) &= 0 \\ \boldsymbol{\mu}_{\bar{\mathbb{A}}} &= 0\end{aligned}$$

and $\mathbf{h}_{\bar{\mathbb{A}}}(\mathbf{w},\mathbf{p}) < 0$, $\boldsymbol{\mu}_{\mathbb{A}} > 0$

Let's define:

$$\mathbf{R}(\mathbf{z},\mathbf{p}) = \left[\begin{array}{c}\nabla_{\mathbf{w}}\Phi(\mathbf{w}) + \nabla_{\mathbf{w}}\mathbf{g}(\mathbf{w},\mathbf{p})\boldsymbol{\lambda} + \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\boldsymbol{\mu}_{\mathbb{A}} \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}_{\mathbb{A}}(\mathbf{w},\mathbf{p})\end{array}\right], \quad \mathbf{z} = \left[\begin{array}{c}\mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu}_{\mathbb{A}}\end{array}\right]$$

Sensitivity given by $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$ with:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{z}} = \left[\begin{array}{ccc} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}} \\ \nabla_{\mathbf{w}}\mathbf{g}^{\top} & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}^{\top} & 0 & 0\end{array}\right]$$

In SQP methods, matrix $\frac{\partial \mathbf{R}}{\partial \mathbf{z}}$ inside underlying Active-Set QP solvers. I.e. we get the sensitivities for free† in SQP !!

†however this factorisation may be hidden deep inside a code you don't have access to... ☹☹☹. Lobby for free access to the factorisation of the KKT matrix !!

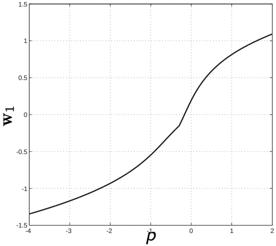
# Linear Predictor

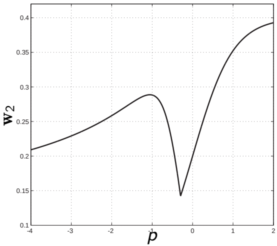
**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

Example

$$\begin{aligned} \min_{\mathbf{w}} \quad & \left\|\mathbf{w}\right\|^2 \\ \text{s.t.} \quad & \mathbf{w}_2 - \mathbf{w}_1\left(1+\mathbf{w}_1^2\right) + p = 0 \\ & \frac{1}{5}\left(\tanh p + 1\right) - \mathbf{w}_2 \leq 0 \end{aligned}$$



# Linear Predictor

**Parametric NLP:**

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

Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

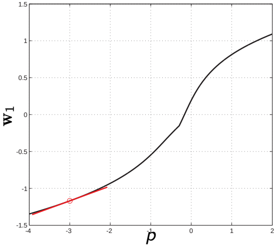
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$

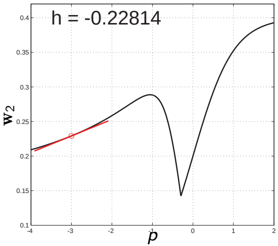
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\begin{aligned} \min_{\mathbf{w}} \quad & \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0 \\ & \frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0 \end{aligned}$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

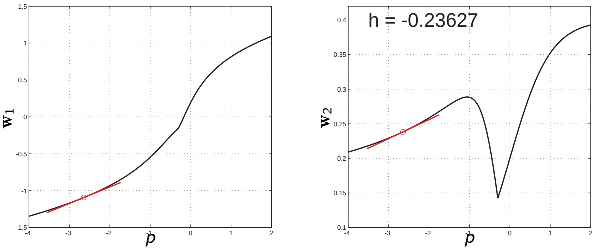
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

$$\frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1+\mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

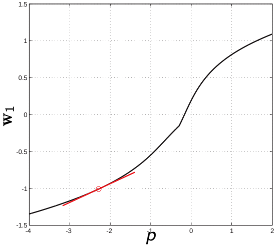
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$

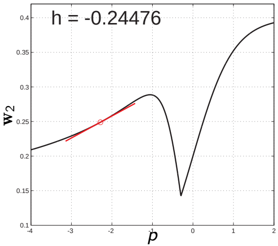
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

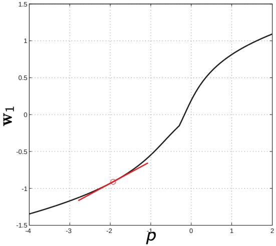
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

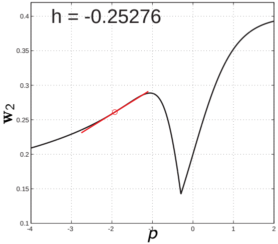
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1+\mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

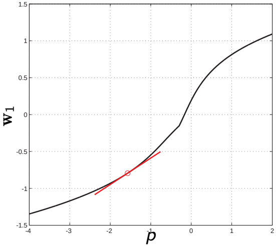
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

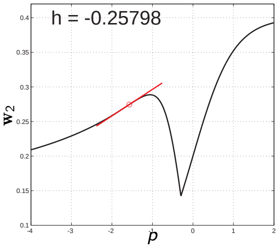
$$\frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}\left(\tanh p + 1\right) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

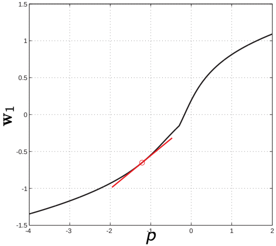
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

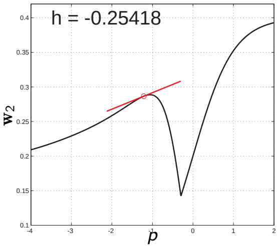
$$\frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1+\mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

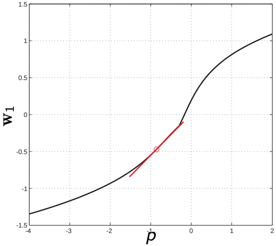
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

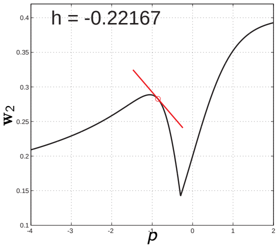
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1+\mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

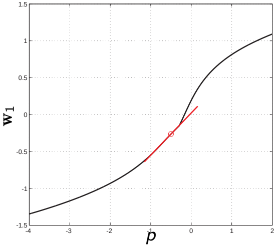
Constraint $\mathbf{h}$ inactive, $\boldsymbol{\mu} = 0$, $\mathbb{A} = \varnothing$ and:

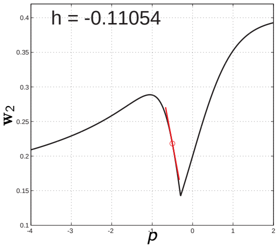
$$\frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \end{bmatrix}$$
$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{p}} = 0$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}\left(\tanh p + 1\right) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

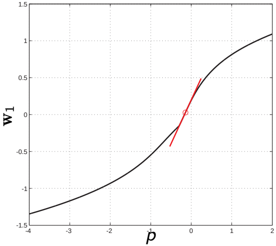
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}} \mathbf{g} & \nabla_{\mathbf{w}} \mathbf{h} \\ \nabla_{\mathbf{w}} \mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}} \mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

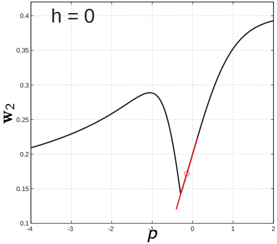
Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1 \left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$

**Change of Active Set!!**



# Linear Predictor

**Parametric NLP:**

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

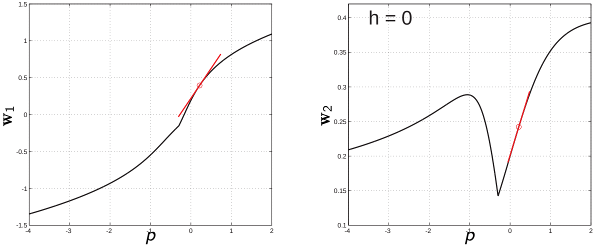
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}} \mathbf{g} & \nabla_{\mathbf{w}} \mathbf{h} \\ \nabla_{\mathbf{w}} \mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}} \mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\begin{aligned} \min_{\mathbf{w}} \quad & \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0 \\ & \frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0 \end{aligned}$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

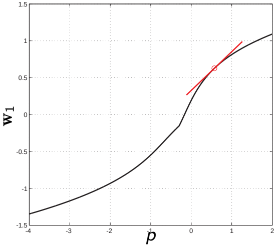
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

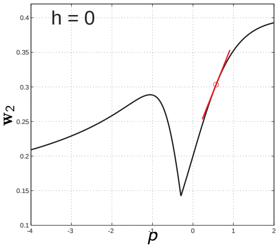
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}} \mathbf{g} & \nabla_{\mathbf{w}} \mathbf{h} \\ \nabla_{\mathbf{w}} \mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}} \mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

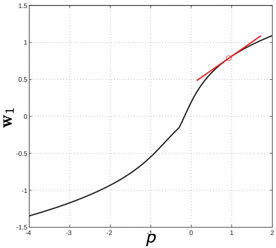
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

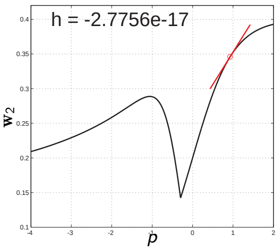
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}} \mathbf{g} & \nabla_{\mathbf{w}} \mathbf{h} \\ \nabla_{\mathbf{w}} \mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}} \mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1 \left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

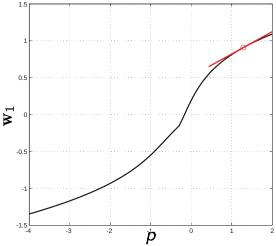
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

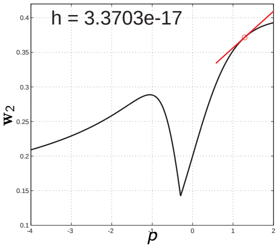
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}} \mathbf{g} & \nabla_{\mathbf{w}} \mathbf{h} \\ \nabla_{\mathbf{w}} \mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}} \mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \le 0$$

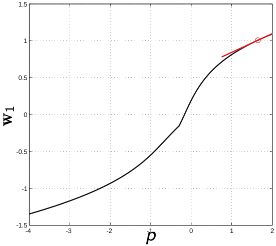
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

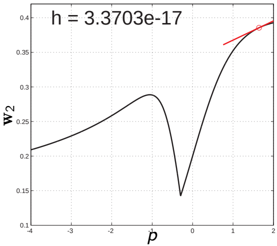
$$\frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}}\begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \le 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$



# Linear Predictor

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

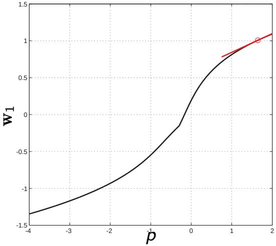
Constraint $\mathbf{h}$ active, $\boldsymbol{\mu} \neq 0$, $\mathbb{A} = \{1\}$ and:

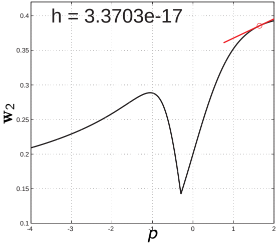
$$\frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = - \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}^\top & 0 & 0 \end{bmatrix}^{-1} \frac{\partial}{\partial \mathbf{p}} \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} \end{bmatrix}$$

Example

$$\min_{\mathbf{w}} \quad \|\mathbf{w}\|^2$$
$$\text{s.t.} \quad \mathbf{w}_2 - \mathbf{w}_1\left(1 + \mathbf{w}_1^2\right) + p = 0$$
$$\frac{1}{5}(\tanh p + 1) - \mathbf{w}_2 \leq 0$$

Check $\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$




At the "corner", derivative does not exist. But **directional derivatives** do !! I.e.

$$\lim_{\mathbf{p}\to\mathbf{p}_-} \frac{\partial \mathbf{z}}{\partial \mathbf{p}} \quad \text{and} \quad \lim_{\mathbf{p}\to\mathbf{p}_+} \frac{\partial \mathbf{z}}{\partial \mathbf{p}} \quad \text{exist}$$

# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

**Linear predictor**

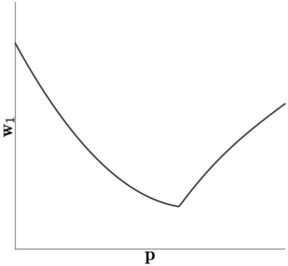
$$\begin{bmatrix} \Delta\mathbf{w} \\ \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{\mu}_{\mathbb{A}} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}^\top & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla_{\mathbf{wp}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \\ \nabla_{\mathbf{p}}\mathbf{h}_{\mathbb{A}}^\top \end{bmatrix} \Delta\mathbf{p}$$

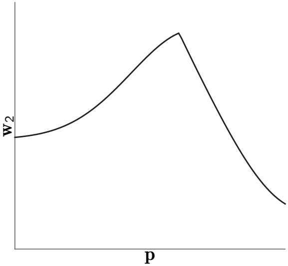
$$\Delta\boldsymbol{\mu}_{\bar{\mathbb{A}}} = 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Linear predictor**

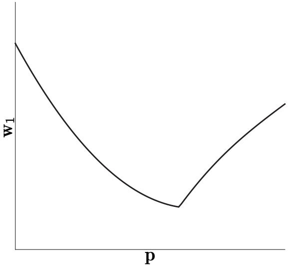
$$\begin{bmatrix} \Delta\mathbf{w} \\ \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{\mu}_{\mathbb{A}} \end{bmatrix} = -\begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}} \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}_{\mathbb{A}}^\top & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla_{\mathbf{w}\mathbf{p}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \\ \nabla_{\mathbf{p}}\mathbf{h}_{\mathbb{A}}^\top \end{bmatrix} \Delta\mathbf{p}$$

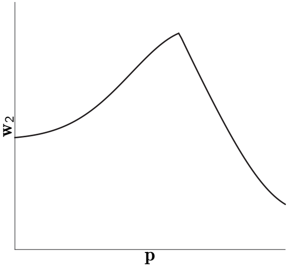
$$\Delta\boldsymbol{\mu}_{\bar{\mathbb{A}}} = 0$$

**How to see the corner?!?**

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

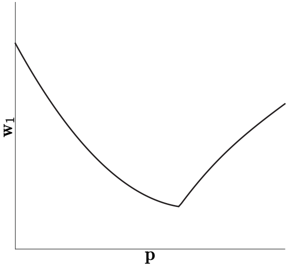
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

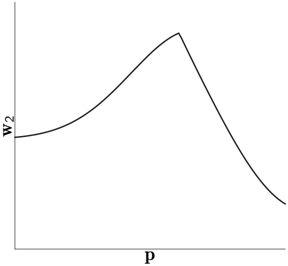
**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

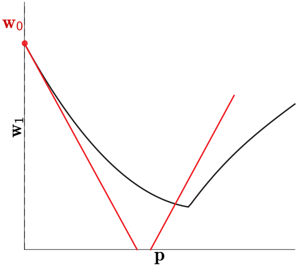
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

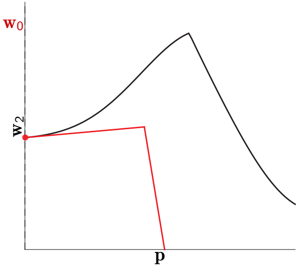
**Non-smooth "linear" predictor**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0 \end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

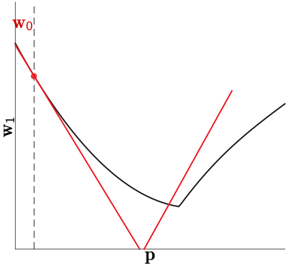
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

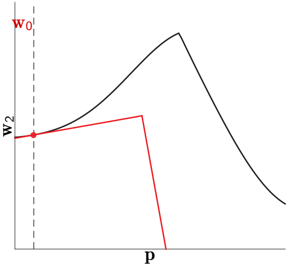
**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

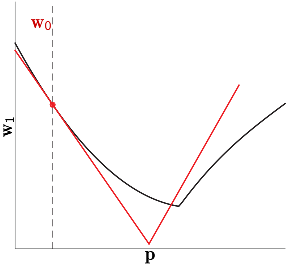
**Non-smooth "linear" predictor**

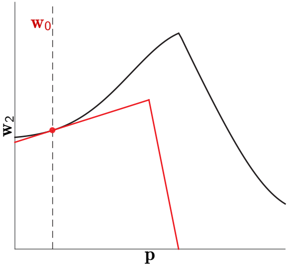
$$\begin{aligned}\min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0\end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

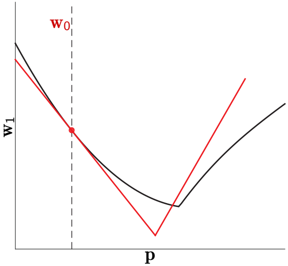
**Non-smooth "linear" predictor**

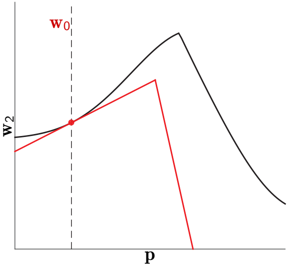
$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

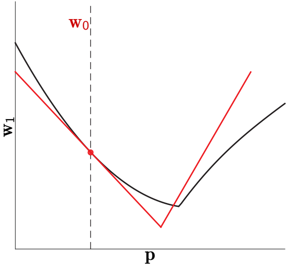
**Non-smooth "linear" predictor**

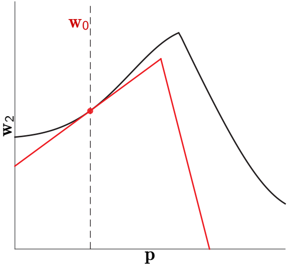
$$\begin{aligned} \min_{\Delta \mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0 \end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

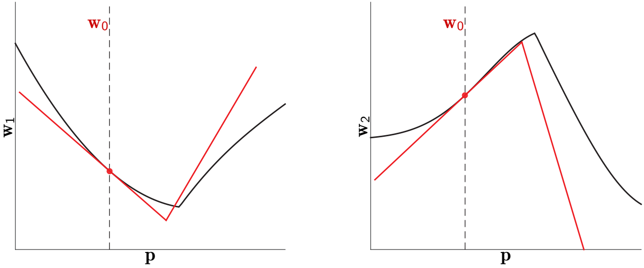
**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Non-smooth "linear" predictor**

$$\min_{\Delta \mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0$$

where




$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$

# The non-smooth "linear" predictor - QP approximation

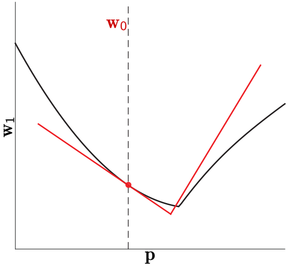
**Parametric NLP:**

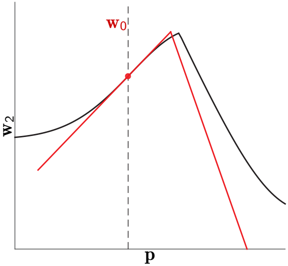
$$\begin{aligned}\min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0\end{aligned}$$

**Non-smooth "linear" predictor**

$$\begin{aligned}\min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0\end{aligned}$$

where




$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$

# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

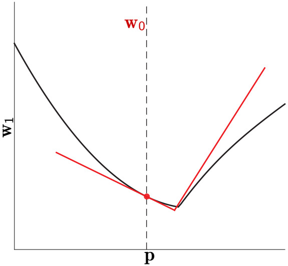
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

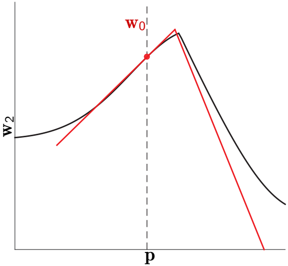
**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

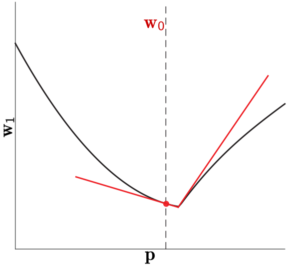
**Parametric NLP:**

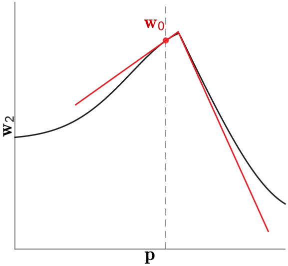
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where




$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$

# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

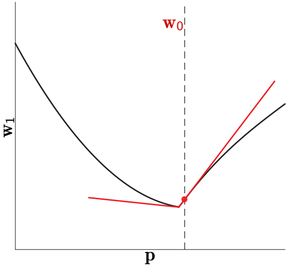
**Non-smooth "linear" predictor**

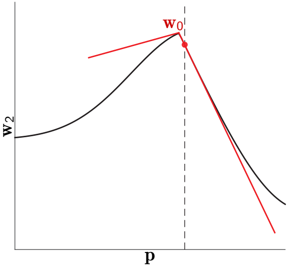
$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}} \mathcal{L} \Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0 \end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

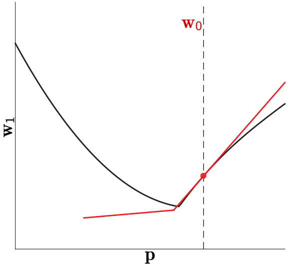
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

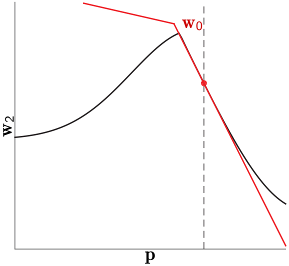
**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

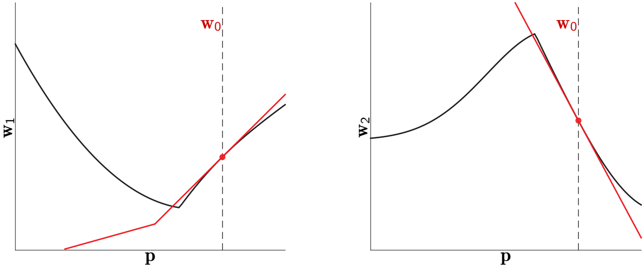
**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

where




$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$

# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

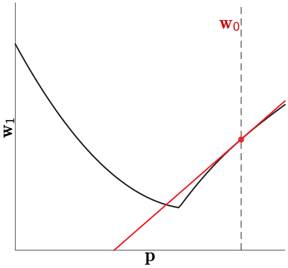
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

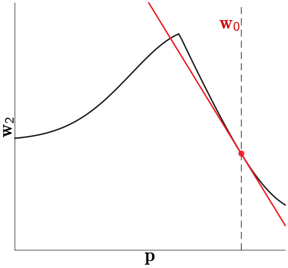
**Non-smooth "linear" predictor**

$$\min_{\Delta \mathbf{w}} \quad \frac{1}{2}\Delta \mathbf{w}^\top H \Delta \mathbf{w} + \Delta \mathbf{p}^\top \nabla_{\mathbf{pw}} \mathcal{L} \Delta \mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}} \mathbf{g}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{g}^\top \Delta \mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}} \mathbf{h}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{h}^\top \Delta \mathbf{p} \leq 0$$

where

$$\Delta \mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta \mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

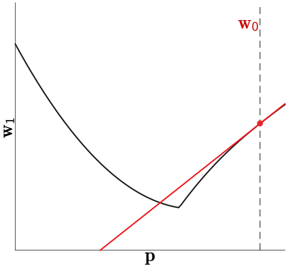
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \le 0$$

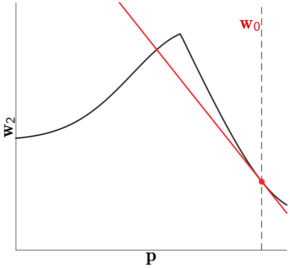
**Non-smooth "linear" predictor**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \le 0$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$
$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



# The non-smooth "linear" predictor - QP approximation

**Parametric NLP:**

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

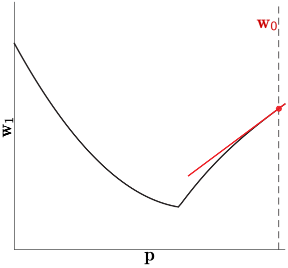
**Non-smooth "linear" predictor**

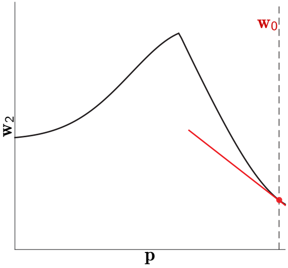
$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

where

$$\Delta\mathbf{w} = \mathbf{w} - \mathbf{w}_0$$

$$\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$$



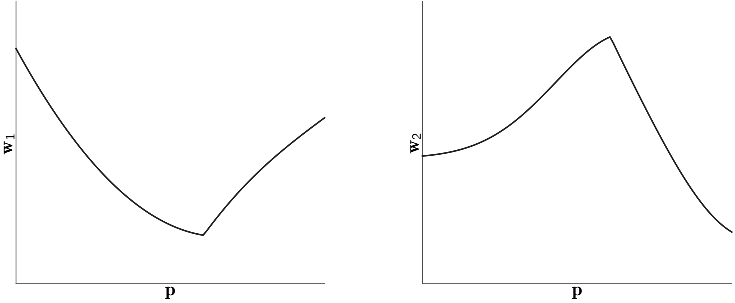

QP predictor holds the sensitivities implicitly (current active set), but also catches a linear approximation of the "kinks" resulting from changes of active set

# The non-smooth "linear" predictor - Implementation

**Parametric NLP:**

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

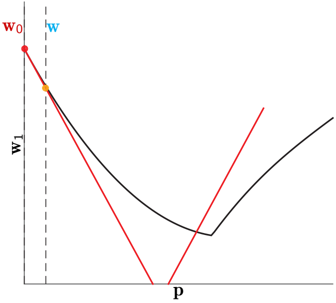
**Parametric NLP:**

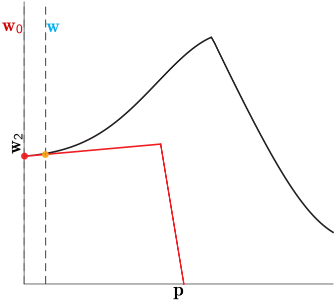
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta \mathbf{w}} \quad \frac{1}{2}\Delta \mathbf{w}^\top H \Delta \mathbf{w} + \Delta \mathbf{p}^\top \nabla_{\mathbf{pw}} \mathcal{L} \Delta \mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}} \mathbf{g}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{g}^\top \Delta \mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}} \mathbf{h}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{h}^\top \Delta \mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta \mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta \mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

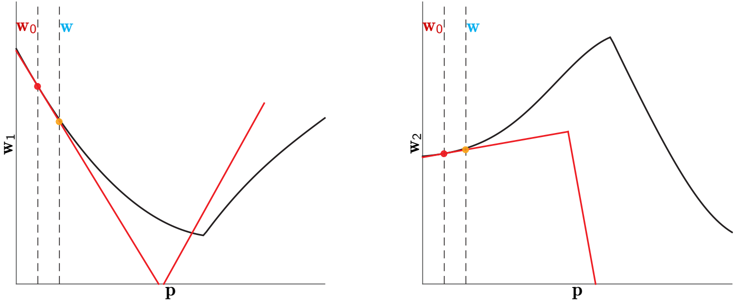
**Parametric NLP:**

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

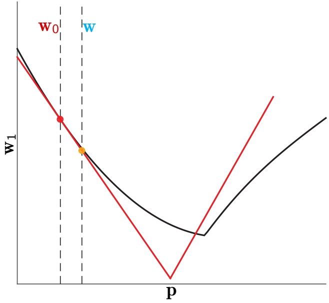
**Parametric NLP:**

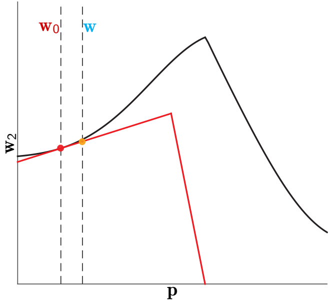
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

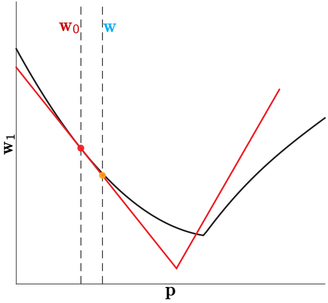
**Parametric NLP:**

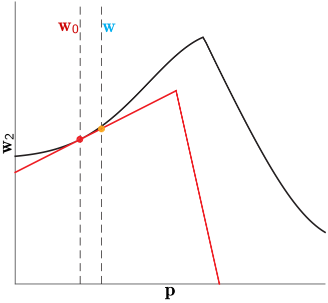
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}, \mathbf{p}) \\ & \mathbf{g}(\mathbf{w}, \mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0 \end{aligned}$$

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

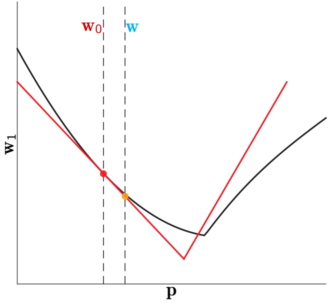
**Parametric NLP:**

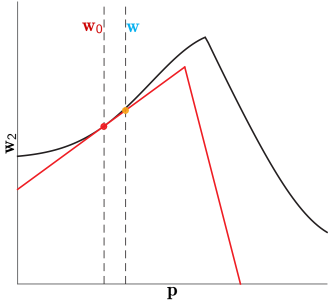
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

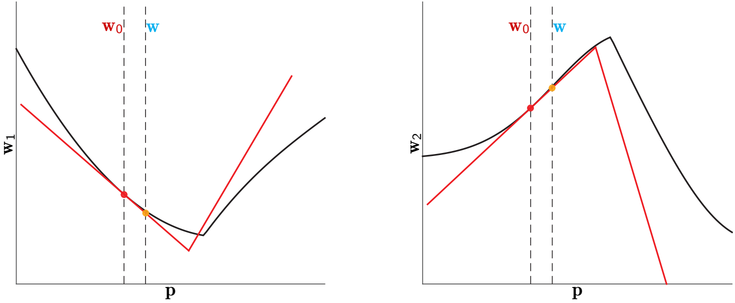
**Parametric NLP:**

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

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

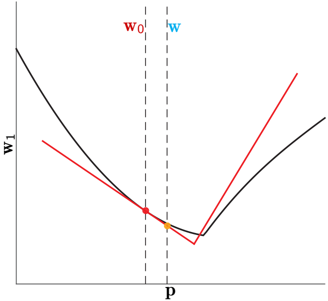
**Parametric NLP:**

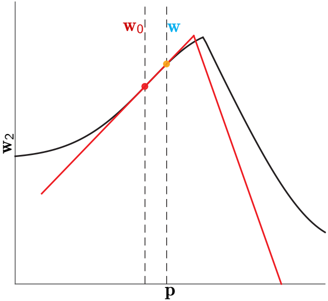
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



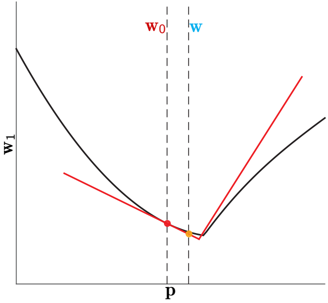
**Parametric NLP:**

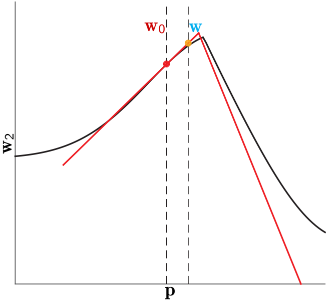
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

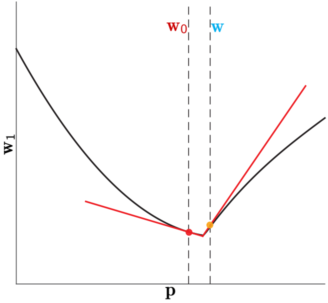
**Parametric NLP:**

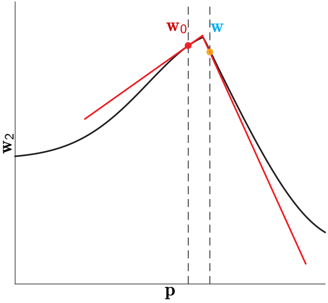
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w},\mathbf{p}\right) \\ & \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) \leq 0 \end{aligned}$$

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H\Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top\Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top\Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top\Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

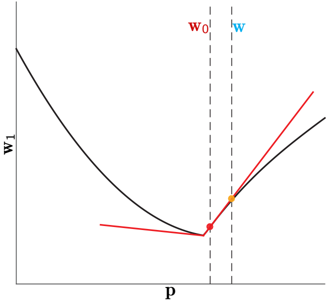
**Parametric NLP:**

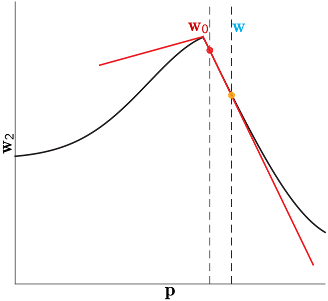
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w},\mathbf{p})$$
$$\mathbf{g}(\mathbf{w},\mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

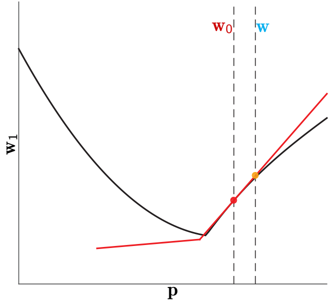
**Parametric NLP:**

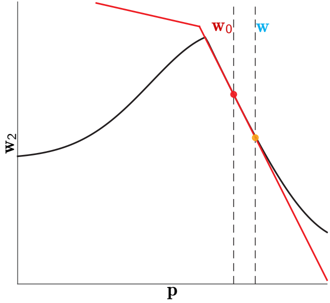
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

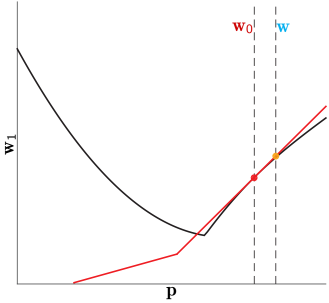
**Parametric NLP:**

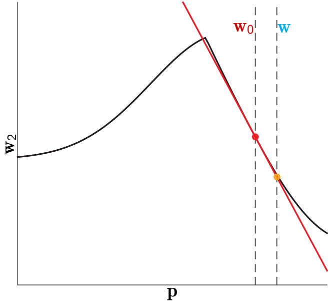
$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w},\mathbf{p}) \\ & \mathbf{g}(\mathbf{w},\mathbf{p}) = 0 \\ & \mathbf{h}(\mathbf{w},\mathbf{p}) \leq 0 \end{aligned}$$

**Approximating QP**

$$\begin{aligned} \min_{\Delta\mathbf{w}} \quad & \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w} \\ & \mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0 \\ & \mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0 \end{aligned}$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# The non-smooth "linear" predictor - Implementation

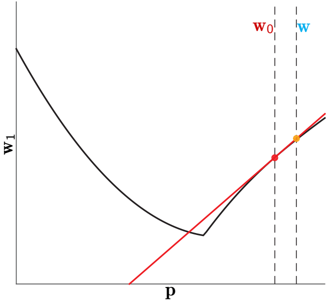
**Parametric NLP:**

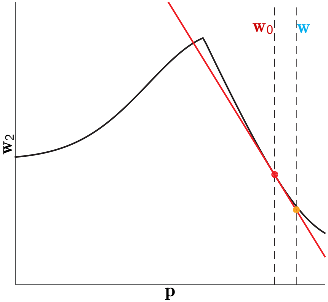
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta \mathbf{w}} \quad \frac{1}{2}\Delta \mathbf{w}^\top H \Delta \mathbf{w} + \Delta \mathbf{p}^\top \nabla_{\mathbf{pw}} \mathcal{L} \Delta \mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}} \mathbf{g}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{g}^\top \Delta \mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}} \mathbf{h}^\top \Delta \mathbf{w} + \nabla_{\mathbf{p}} \mathbf{h}^\top \Delta \mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta \mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta \mathbf{w}$ from the QP



## The non-smooth "linear" predictor - Implementation

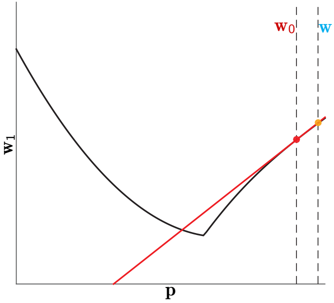
**Parametric NLP:**

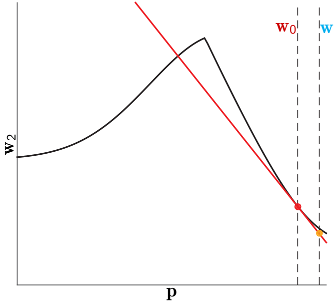
$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w}, \mathbf{p})$$
$$\mathbf{g}(\mathbf{w}, \mathbf{p}) = 0$$
$$\mathbf{h}(\mathbf{w}, \mathbf{p}) \leq 0$$

**Approximating QP**

$$\min_{\Delta\mathbf{w}} \quad \frac{1}{2}\Delta\mathbf{w}^\top H \Delta\mathbf{w} + \Delta\mathbf{p}^\top \nabla_{\mathbf{pw}}\mathcal{L}\Delta\mathbf{w}$$
$$\mathbf{g} + \nabla_{\mathbf{w}}\mathbf{g}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{g}^\top \Delta\mathbf{p} = 0$$
$$\mathbf{h} + \nabla_{\mathbf{w}}\mathbf{h}^\top \Delta\mathbf{w} + \nabla_{\mathbf{p}}\mathbf{h}^\top \Delta\mathbf{p} \leq 0$$

- solved at $\mathbf{p}_0$ yields $\mathbf{w}_0$
- new parameter $\mathbf{p}$, perturbation $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ inserted in the QP
- predicted solution $\mathbf{w} = \mathbf{w}_0 + \Delta\mathbf{w}$ from the QP



# Outline

# Solution Manifolds in IP methods

KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= 0 \\ \mathbf{s} \geq 0, \quad \boldsymbol{\mu} &\geq 0 \end{aligned}$$

IP-KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= \tau \\ \mathbf{s} > 0, \quad \boldsymbol{\mu} &> 0 \end{aligned}$$

# Solution Manifolds in IP methods

KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}} \mathcal{L}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w}, \mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w}, \mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s} \cdot \boldsymbol{\mu} &= 0 \\ \mathbf{s} \geq 0, \quad \boldsymbol{\mu} &\geq 0 \end{aligned}$$

IP-KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}} \mathcal{L}\left(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w}, \mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w}, \mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s} \cdot \boldsymbol{\mu} &= \tau \\ \mathbf{s} > 0, \quad \boldsymbol{\mu} &> 0 \end{aligned}$$

$\tau$ **"smoothens" the corner in** $\mathbf{s} \cdot \boldsymbol{\mu} = 0$**. How does it impact the solution manifold?**

## Solution Manifolds in IP methods

KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= 0 \\ \mathbf{s} \geq 0, \quad \boldsymbol{\mu} &\geq 0 \end{aligned}$$

IP-KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) &= 0 \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) &= 0 \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= \tau \\ \mathbf{s} > 0, \quad \boldsymbol{\mu} &> 0 \end{aligned}$$

$\tau$ **"smoothens" the corner in** $\mathbf{s}\cdot\boldsymbol{\mu} = 0$**. How does it impact the solution manifold?**

- According to $\mathcal{O}\left(\tau\right)$

# Solution Manifolds in IP methods

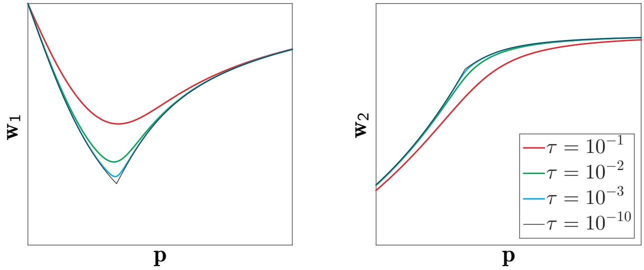
KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= 0 \\ \mathbf{s}\geq 0, \quad \boldsymbol{\mu} &\geq 0 \end{aligned}$$

IP-KKT conditions

$$\begin{aligned} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) &= 0 \\ \mathbf{g}(\mathbf{w},\mathbf{p}) &= 0 \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} &= 0 \\ \mathbf{s}\cdot\boldsymbol{\mu} &= \tau \\ \mathbf{s}> 0, \quad \boldsymbol{\mu} &> 0 \end{aligned}$$

$\tau$ **"smoothens" the corner in $\mathbf{s}\cdot\boldsymbol{\mu}=0$. How does it impact the solution manifold?**




- According to $\mathcal{O}(\tau)$
- Smoothens the corners as well

# Sensitivities in IP methods

IP-KKT conditions

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

where $\mathbf{z} = \{\mathbf{w},\ \boldsymbol{\lambda},\ \boldsymbol{\mu},\ \mathbf{s}\}$

## Sensitivities in IP methods

IP-KKT conditions

$$\mathbf{R}(\mathbf{z}, \mathbf{p}, \tau) = \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{p}) \\ \mathbf{g}(\mathbf{w}, \mathbf{p}) \\ \mathbf{h}(\mathbf{w}, \mathbf{p}) + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

where $\mathbf{z} = \{\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}\}$

Implicit function theorem:

$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\mathbf{p}} = \frac{\partial \mathbf{R}}{\partial \mathbf{z}} \frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{R}}{\partial \mathbf{p}} = 0$$

hence

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Sensitivities in IP methods

IP-KKT conditions

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

where $\mathbf{z} = \{\mathbf{w},\ \boldsymbol{\lambda},\ \boldsymbol{\mu},\ \mathbf{s}\}$

Implicit function theorem:

$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\mathbf{p}} = \frac{\partial \mathbf{R}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{R}}{\partial \mathbf{p}} = 0$$

hence

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

and

$$\frac{\partial \mathbf{R}}{\partial \mathbf{z}} = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h} & 0 \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}^\top & 0 & 0 & I \\ 0 & 0 & \mathrm{diag}(\mathbf{s}) & \mathrm{diag}(\boldsymbol{\mu}) \end{bmatrix}, \qquad \frac{\partial \mathbf{R}}{\partial \mathbf{p}} = \begin{bmatrix} \nabla_{\mathbf{wp}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \\ \nabla_{\mathbf{p}}\mathbf{h}^\top \\ 0 \end{bmatrix}$$

Note: factorization $\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}$ is available from solver (used for solving $\mathbf{R} = 0$)

## Sensitivities in IP methods

IP-KKT conditions

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p})+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

where $\mathbf{z} = \{\mathbf{w},\ \boldsymbol{\lambda},\ \boldsymbol{\mu},\ \mathbf{s}\}$

Implicit function theorem:

$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\mathbf{p}} = \frac{\partial\mathbf{R}}{\partial\mathbf{z}}\frac{\partial\mathbf{z}}{\partial\mathbf{p}} + \frac{\partial\mathbf{R}}{\partial\mathbf{p}} = 0$$

hence

$$\frac{\partial\mathbf{z}}{\partial\mathbf{p}} = -\frac{\partial\mathbf{R}}{\partial\mathbf{z}}^{-1}\frac{\partial\mathbf{R}}{\partial\mathbf{p}}$$

and

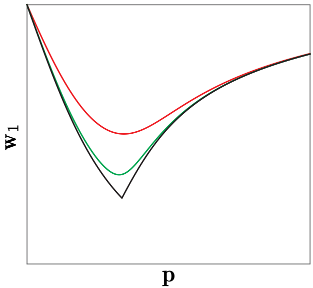
$$\frac{\partial\mathbf{R}}{\partial\mathbf{z}} = \begin{bmatrix} H & \nabla_{\mathbf{w}}\mathbf{g} & \nabla_{\mathbf{w}}\mathbf{h} & 0 \\ \nabla_{\mathbf{w}}\mathbf{g}^\top & 0 & 0 & 0 \\ \nabla_{\mathbf{w}}\mathbf{h}^\top & 0 & 0 & I \\ 0 & 0 & \mathrm{diag}(\mathbf{s}) & \mathrm{diag}(\boldsymbol{\mu}) \end{bmatrix}, \qquad \frac{\partial\mathbf{R}}{\partial\mathbf{p}} = \begin{bmatrix} \nabla_{\mathbf{wp}}\mathcal{L} \\ \nabla_{\mathbf{p}}\mathbf{g}^\top \\ \nabla_{\mathbf{p}}\mathbf{h}^\top \\ 0 \end{bmatrix}$$

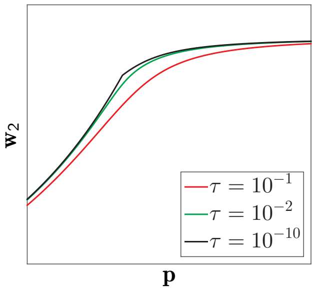
Note: factorization $\frac{\partial\mathbf{R}}{\partial\mathbf{z}}^{-1}$ is available from solver (used for solving $\mathbf{R}=0$)

Linear predictor:

$$\mathbf{z}^\star(\mathbf{p}+\Delta\mathbf{p},\tau) = \mathbf{z}^\star(\mathbf{p},\tau) + \left.\frac{\partial\mathbf{z}}{\partial\mathbf{p}}\right|_{\mathbf{z}^\star(\mathbf{p}),\mathbf{p},\tau}\Delta\mathbf{p}$$

# Solution Manifolds & Sensitivities - Illustration



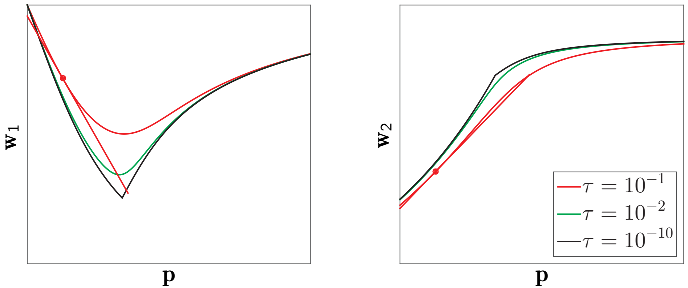

IP-KKT

$$\mathbf{R}(\mathbf{z}, \mathbf{p}, \tau) = \begin{bmatrix} \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{p}) \\ \mathbf{g}(\mathbf{w}, \mathbf{p}) \\ \mathbf{h}(\mathbf{w}, \mathbf{p}) + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



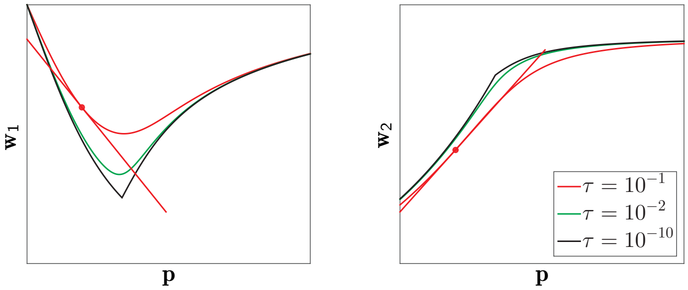

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



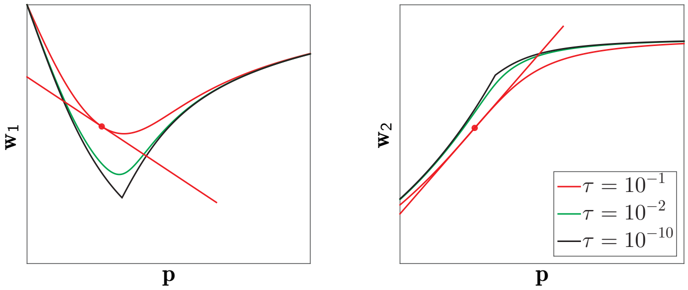

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p})+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



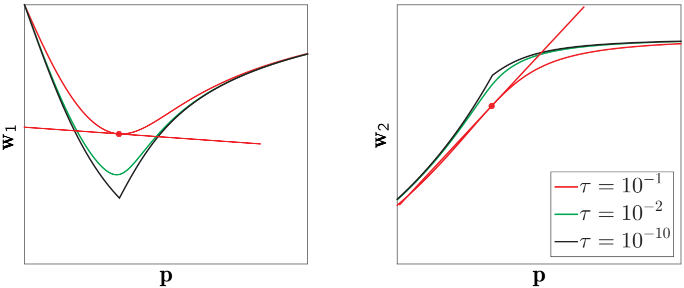

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



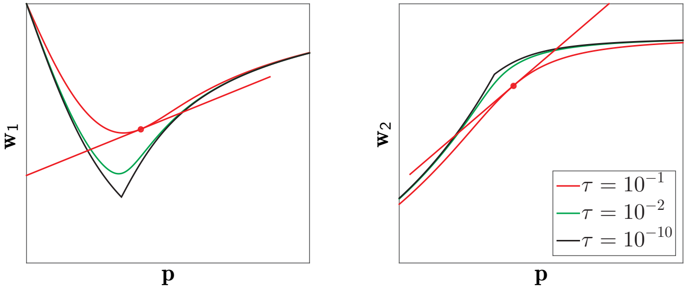

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



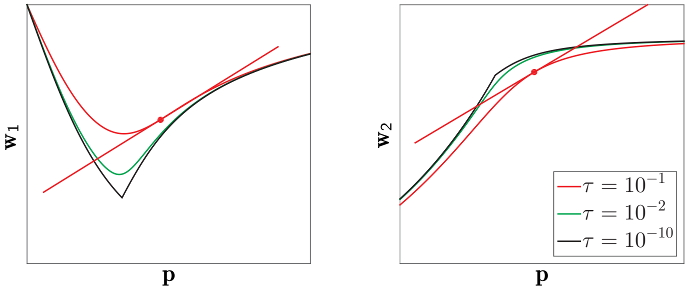

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



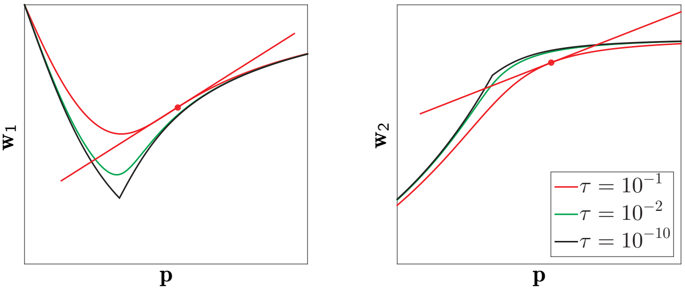

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

## Solution Manifolds & Sensitivities - Illustration



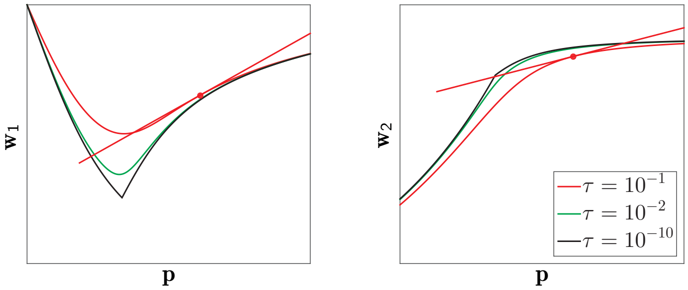

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



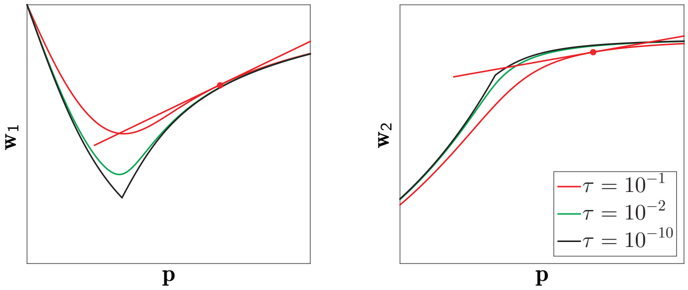

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



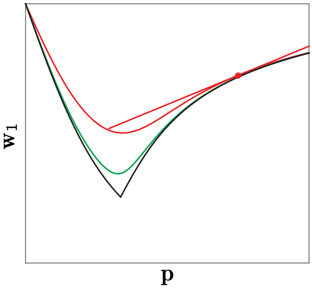

IP-KKT

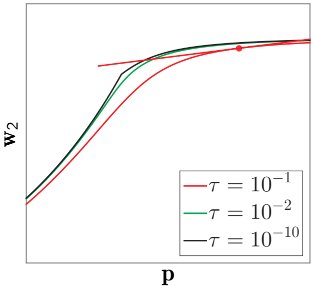
$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



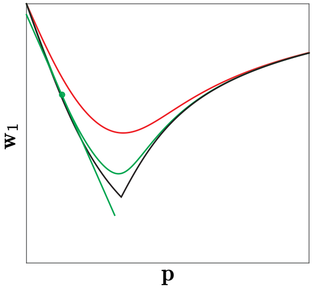

IP-KKT

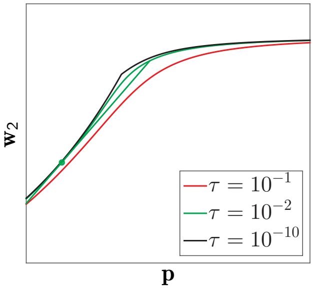
$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



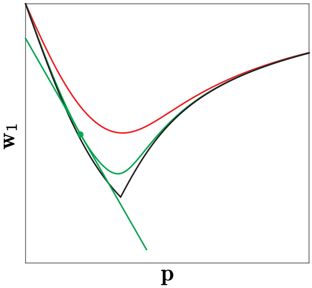

IP-KKT

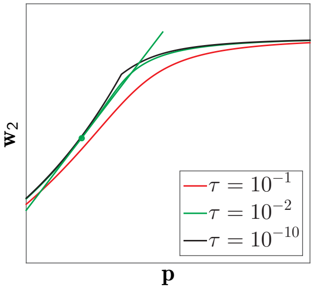
$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



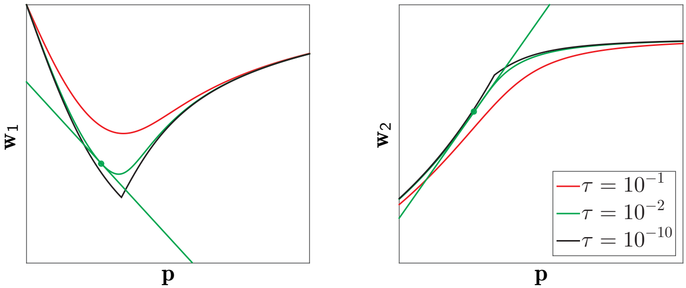

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



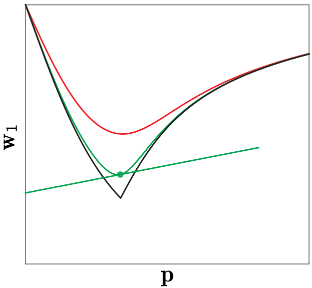

IP-KKT

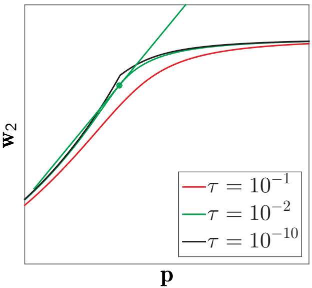
$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



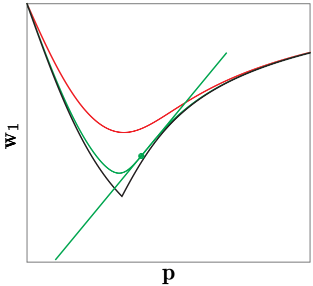

IP-KKT

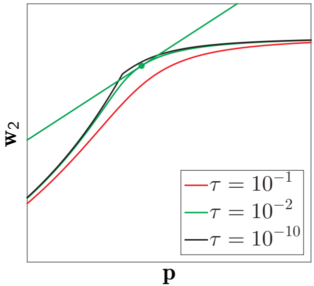
$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



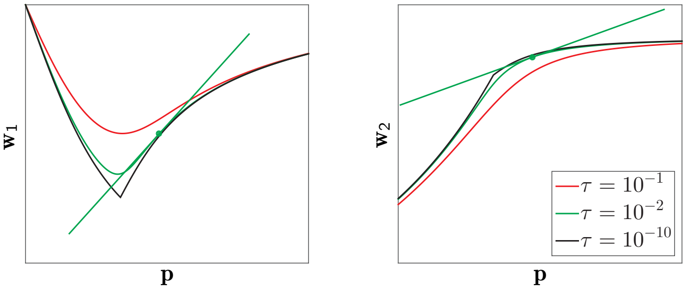

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



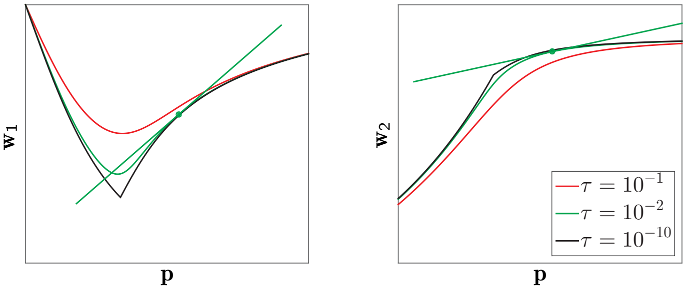

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



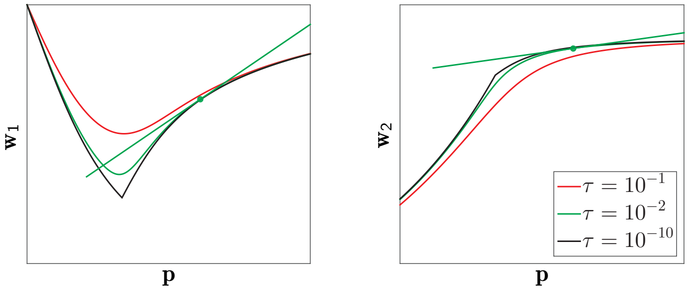

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



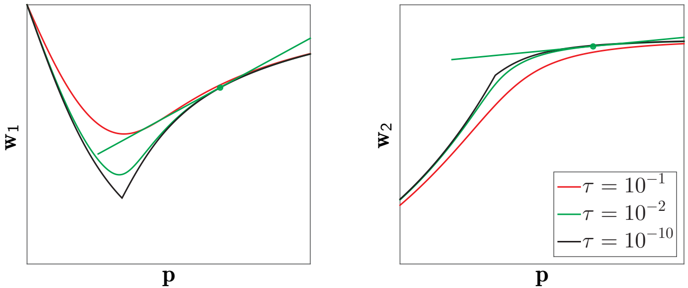

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



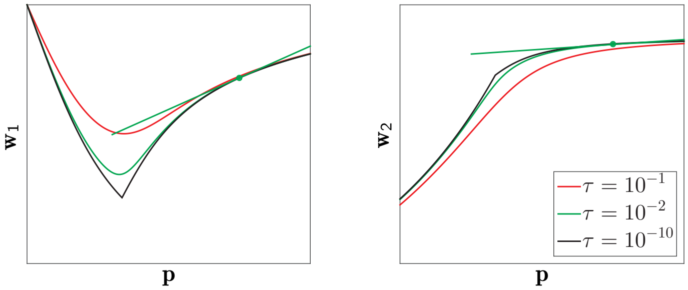

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



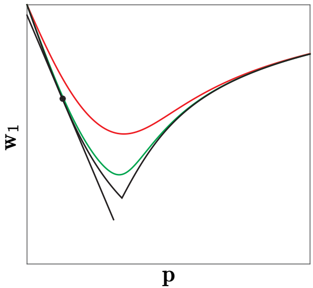

IP-KKT

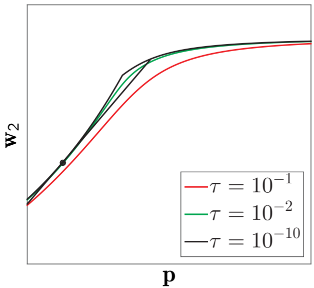
$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



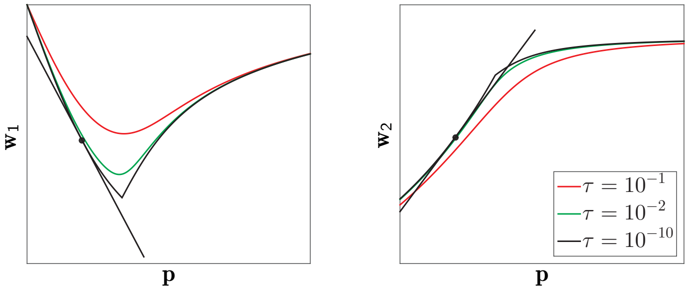

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



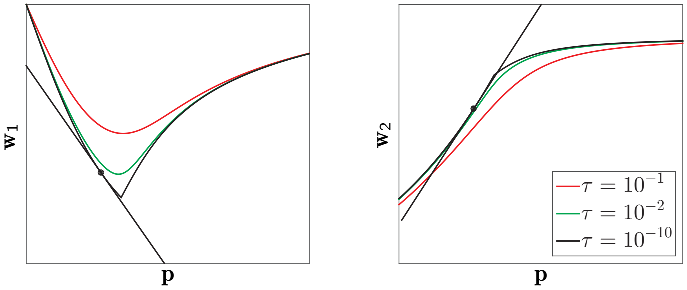

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



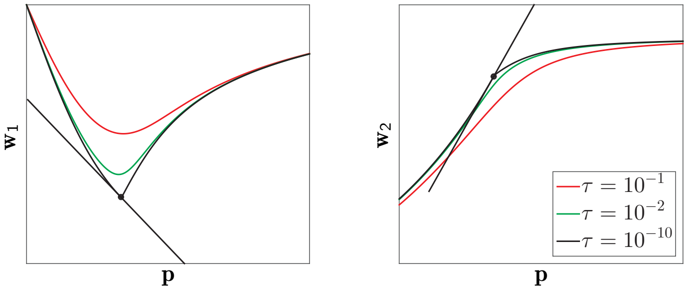

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



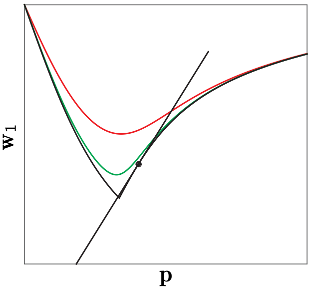

IP-KKT

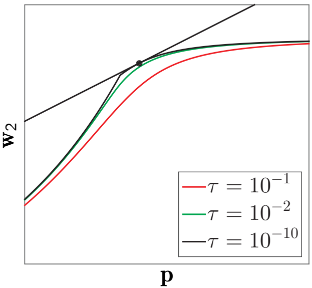
$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



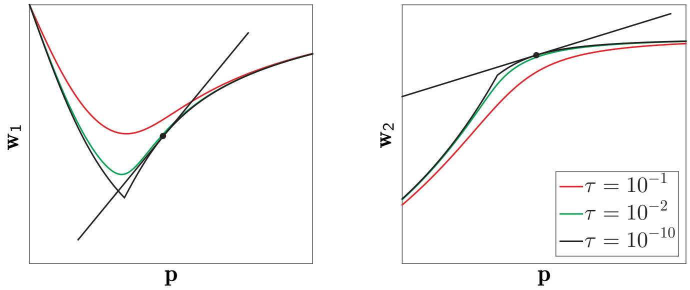

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



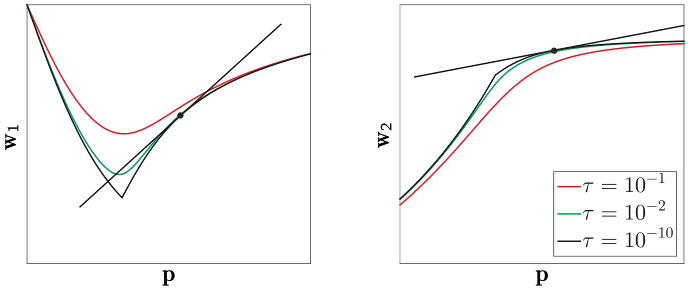

IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



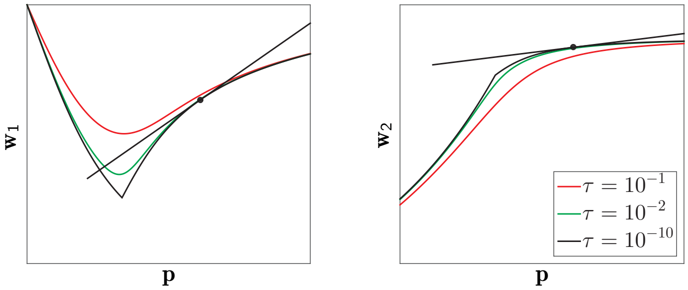

IP-KKT

$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



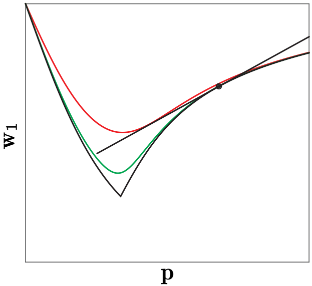

IP-KKT

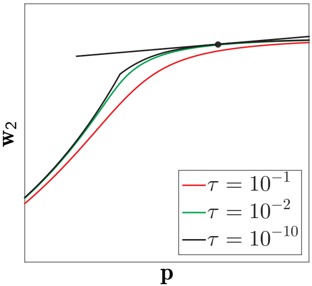
$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration



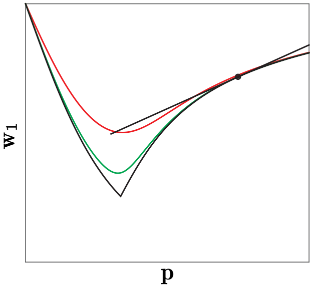

IP-KKT

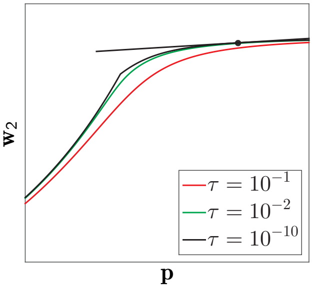
$$\mathbf{R}(\mathbf{z},\mathbf{p},\tau) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}) \\ \mathbf{g}(\mathbf{w},\mathbf{p}) \\ \mathbf{h}(\mathbf{w},\mathbf{p}) + \mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

# Solution Manifolds & Sensitivities - Illustration




IP-KKT

$$\mathbf{R}\left(\mathbf{z},\mathbf{p},\tau\right) = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L}\left(\mathbf{w},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{p}\right) \\ \mathbf{g}\left(\mathbf{w},\mathbf{p}\right) \\ \mathbf{h}\left(\mathbf{w},\mathbf{p}\right)+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

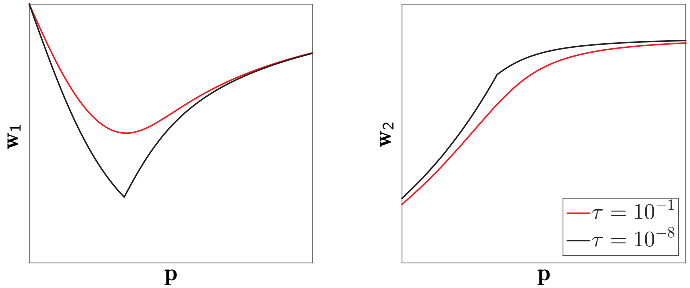
Tangential predictor:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

**Validity of linear predictor becomes limited when going through sharp turns, i.e. AS change & small $\tau$**

# Outline

# Linear Prediction for IP methods

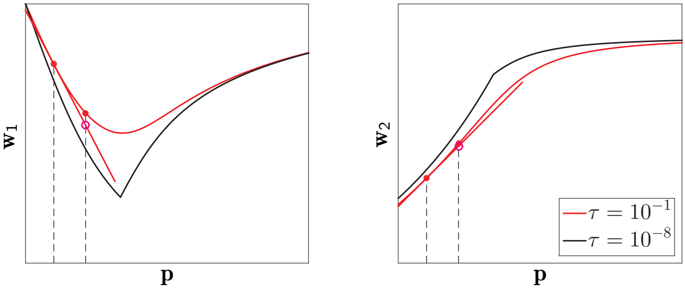



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

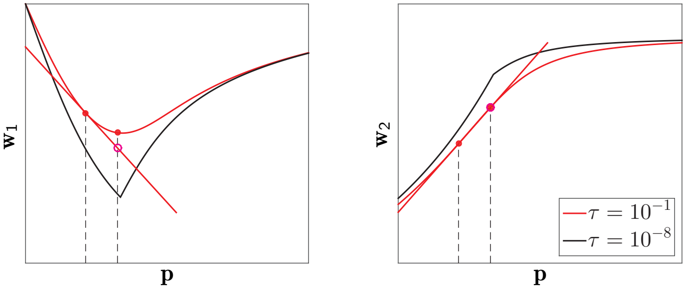



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

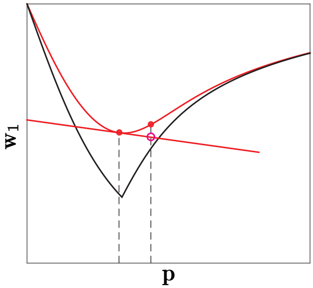


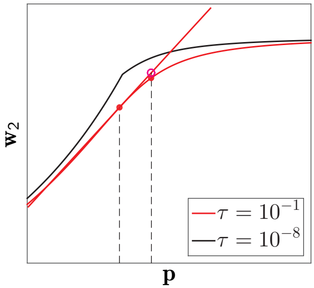

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods



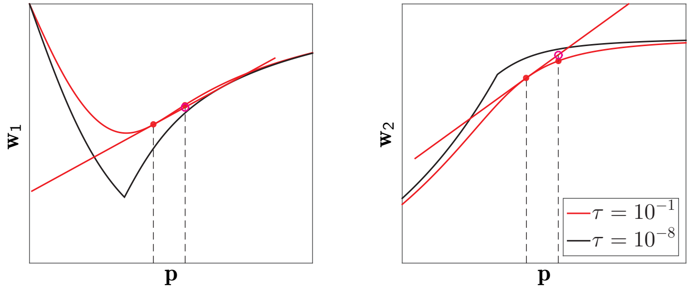




$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

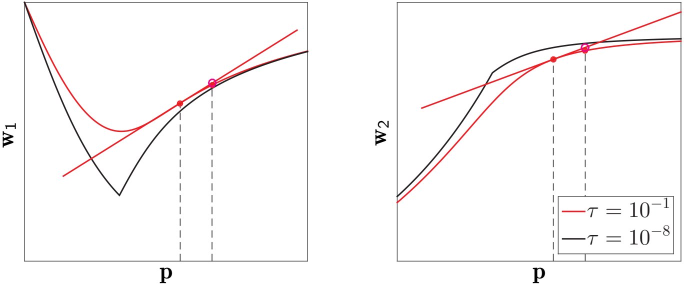



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

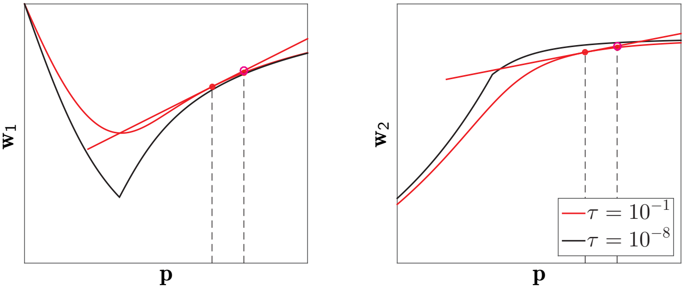



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

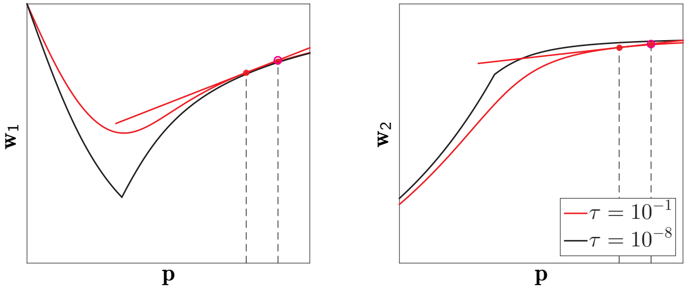



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

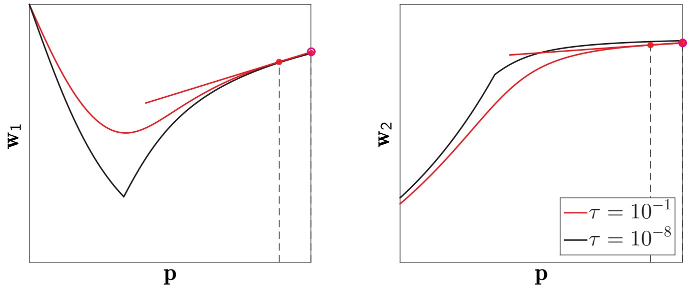



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ “large” : “wrong manifold” but smoother
- Correct manifold at $\tau$ small, but predictor can be “lost” (manifold has “corners”)

# Linear Prediction for IP methods

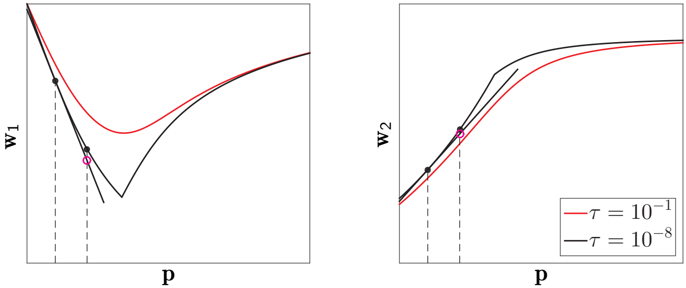



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

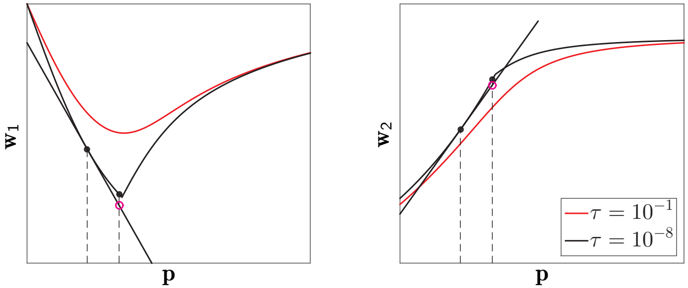



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

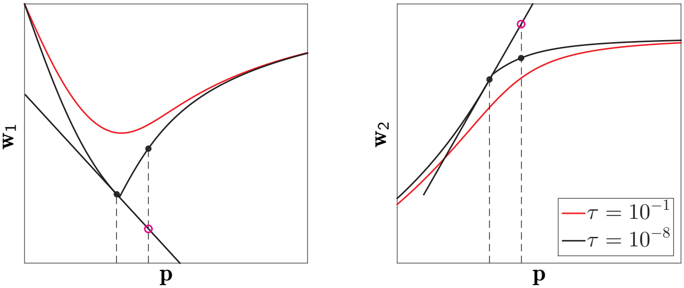



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

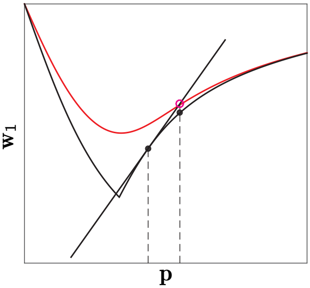


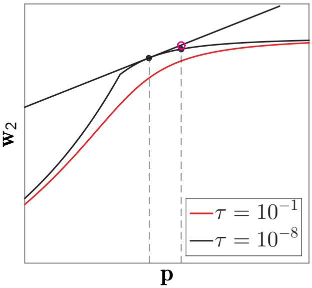

$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods



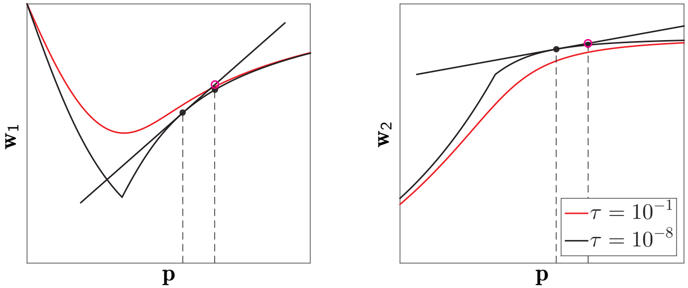




$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

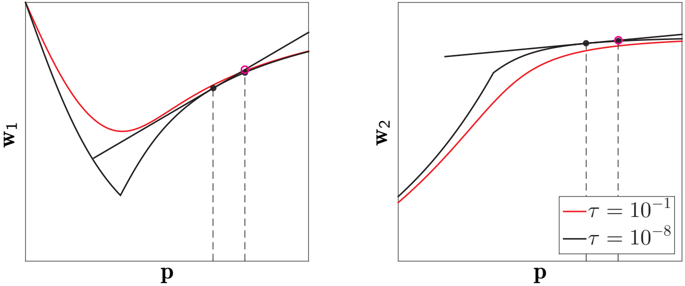



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

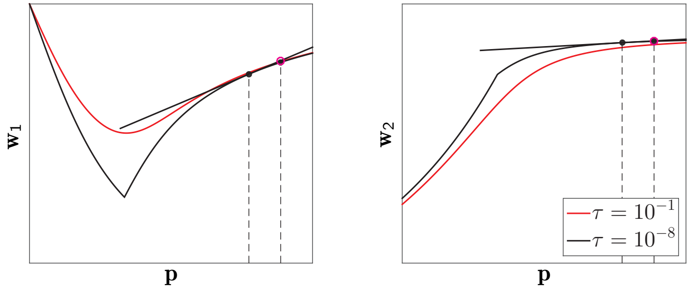



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h} + \mathbf{s} \\ \mathbf{s} \cdot \boldsymbol{\mu} - \tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods

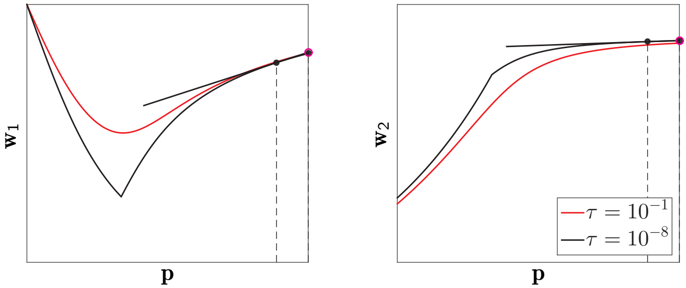



$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Linear Prediction for IP methods




$$\mathbf{R} = \begin{bmatrix} \nabla_{\mathbf{w}}\mathcal{L} \\ \mathbf{g} \\ \mathbf{h}+\mathbf{s} \\ \mathbf{s}\cdot\boldsymbol{\mu}-\tau \end{bmatrix} = 0$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{p}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}$$

- Predictor is effective at $\tau$ "large" : "wrong manifold" but smoother
- Correct manifold at $\tau$ small, but predictor can be "lost" (manifold has "corners")

# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$

# Prediction-Corrector for IP methods

- Predict solution (linear predictor)
$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$
- Correct IP-KKT error
$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

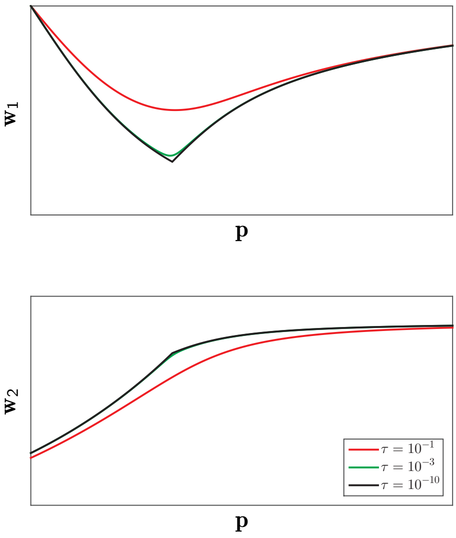
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

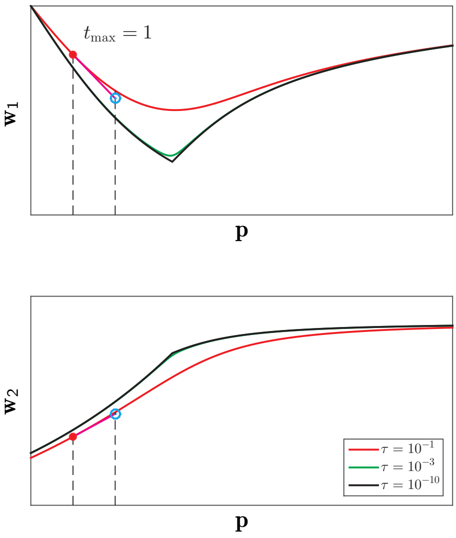
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

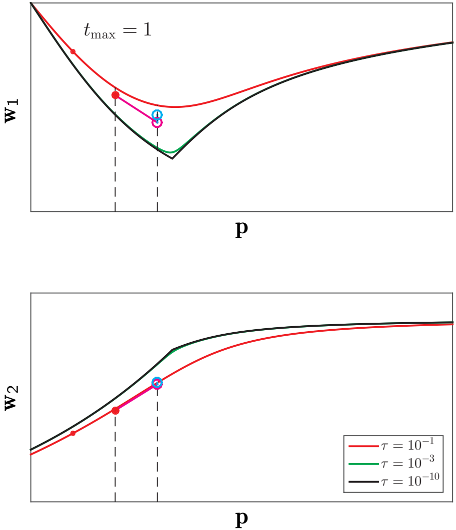
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

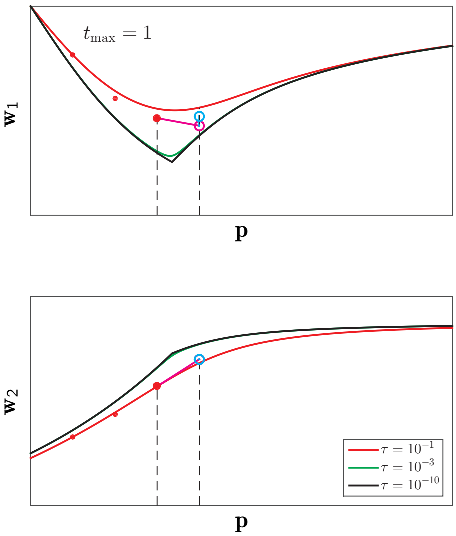
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

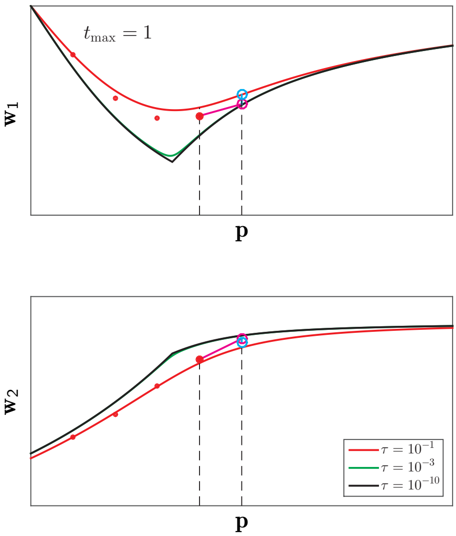
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

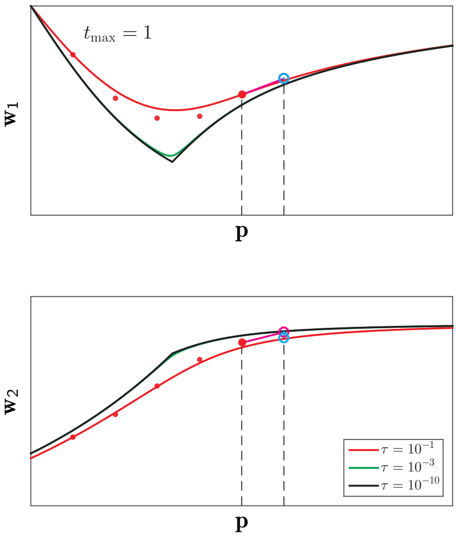
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\text{max}} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\text{max}}\Delta\mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\text{max}}\Delta\boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\text{max}}\Delta\mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

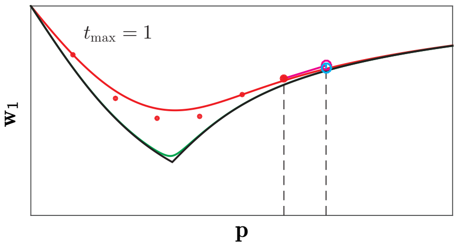
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \bigg|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

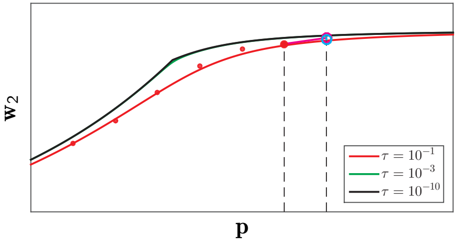
Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

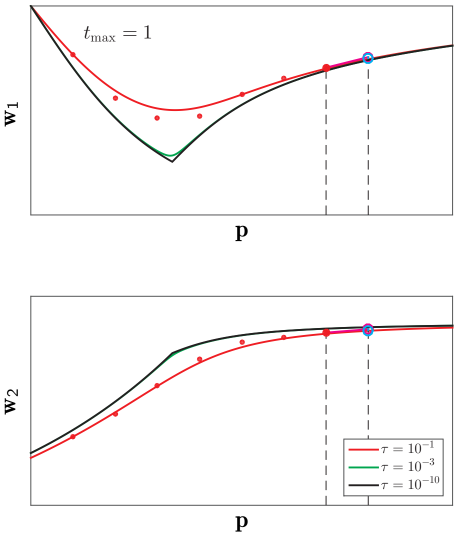
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

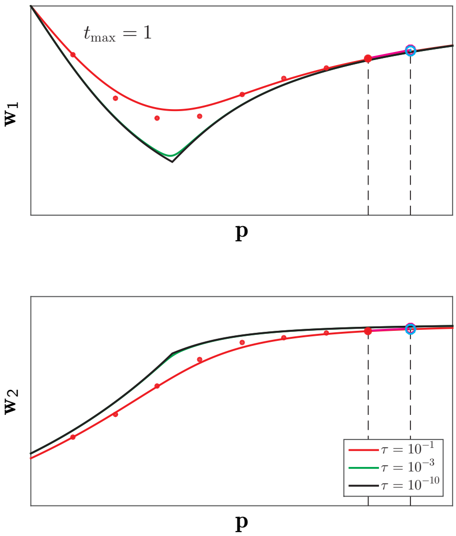
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

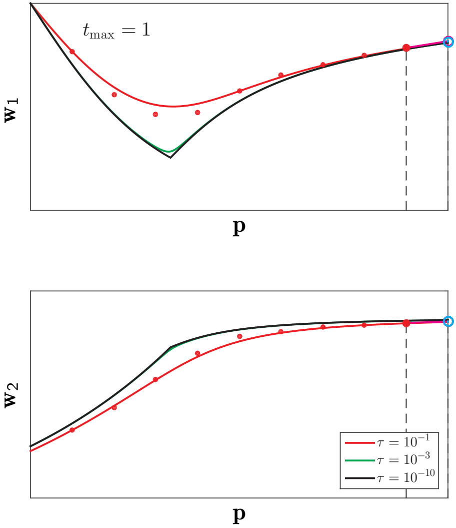
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

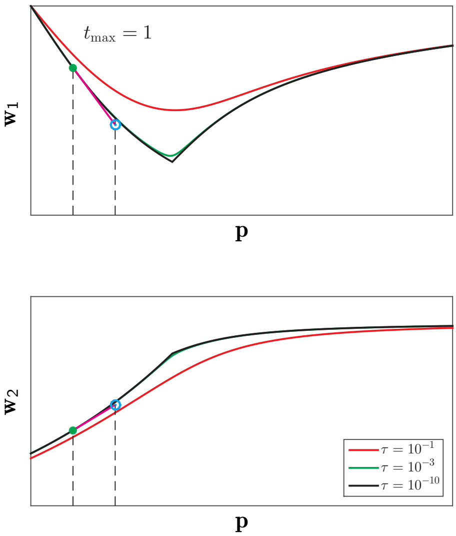
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



## Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

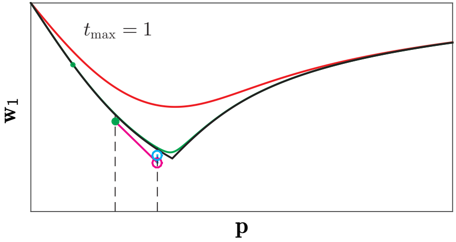
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

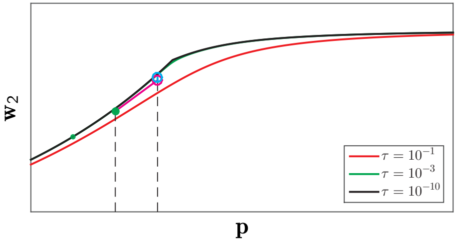
Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

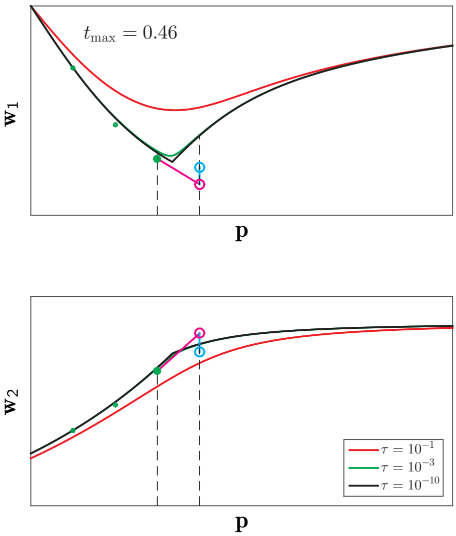
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Bigg|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

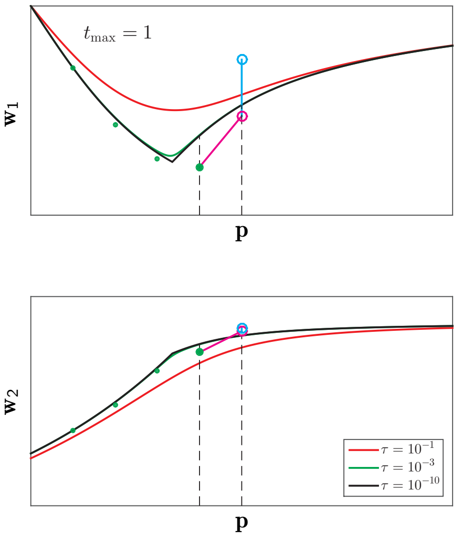
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

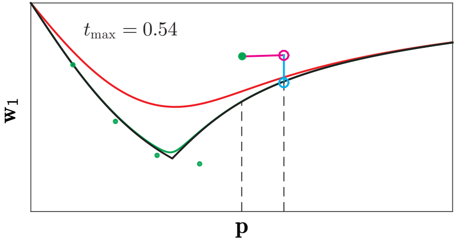
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

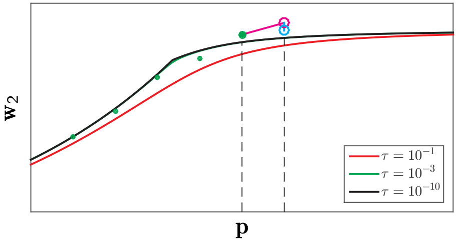
Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

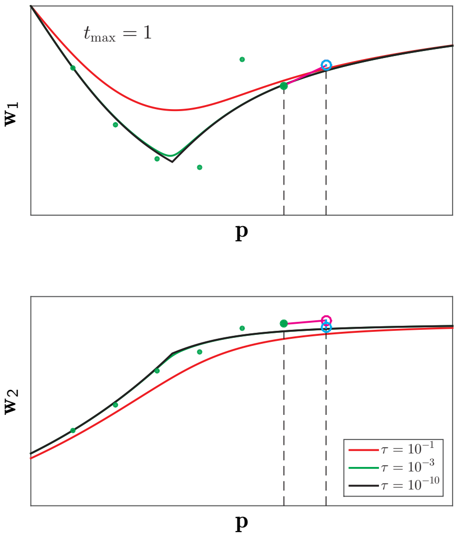
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\bigg|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

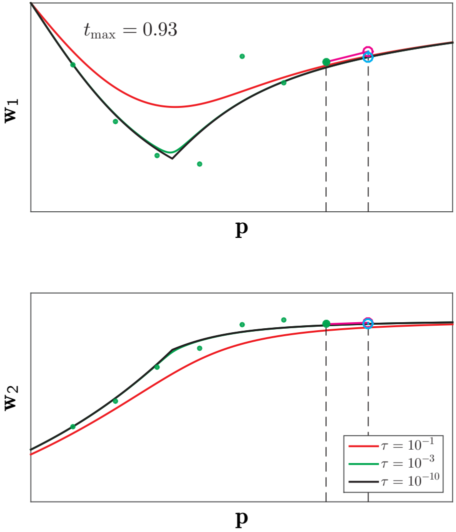
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\text{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\text{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$
Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

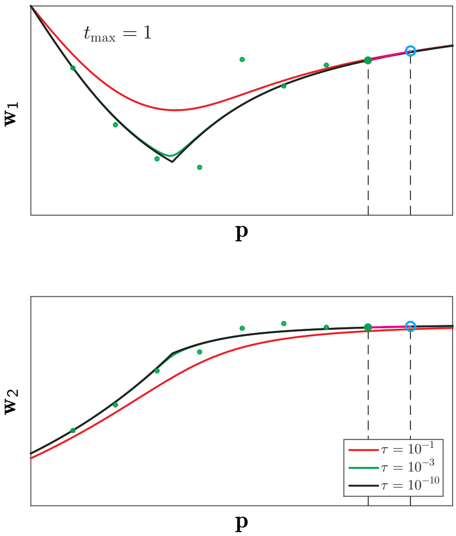
$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \frac{\partial \mathbf{R}}{\partial \mathbf{p}} \Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1} \mathbf{R} \Big|_{\hat{\mathbf{z}}(\mathbf{p}),\, \mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max} \Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max} \Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max} \Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$



# Prediction-Corrector for IP methods

- Predict solution (linear predictor)

$$\Delta \mathbf{z}_{\mathrm{pred}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p}$$

- Correct IP-KKT error

$$\Delta \mathbf{z}_{\mathrm{corr}} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}$$

**Algorithm:** Path-following IP

**Input:** Estimate $\hat{\mathbf{z}}(\mathbf{p})$, parameter $\mathbf{p}_+$

Predictor-corrector: $\Delta \mathbf{p} = \mathbf{p}_+ - \mathbf{p}$

$$\Delta \mathbf{z} = -\frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\frac{\partial \mathbf{R}}{\partial \mathbf{p}}\Delta \mathbf{p} - \frac{\partial \mathbf{R}}{\partial \mathbf{z}}^{-1}\mathbf{R}\Big|_{\hat{\mathbf{z}}(\mathbf{p}),\,\mathbf{p}}$$

Step-size: find $t_{\max} \in ]0, 1]$ such that:

$$\mathbf{s} + t_{\max}\Delta \mathbf{s} > 0, \qquad \boldsymbol{\mu} + t_{\max}\Delta \boldsymbol{\mu} > 0$$

Update $\hat{\mathbf{z}}(\mathbf{p}_+) = \hat{\mathbf{z}}(\mathbf{p}) + t_{\max}\Delta \mathbf{z}$

**return** $\hat{\mathbf{z}}(\mathbf{p}_+)$