# Lecture 2: nonlinear constrained optimization. Null space gradient flows

Florian Feppon

Spring 2022 – Seminar for Applied Mathematics

ETH zürich

# Equality constrained optimization

Consider the optimization problem

$$\min_{x \in V} \quad J(x) \\ \text{s.t. } \boldsymbol{g}(x) = 0 \tag{1}$$

Assume that $\operatorname{rank}(\mathrm{D}\boldsymbol{g}(x)\mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}}) = p$.

## Definition 1

The *null space* and *range space* directions $\boldsymbol{\xi}_J(x)$ and $\boldsymbol{\xi}_C(x)$ are defined by:

$$\boldsymbol{\xi}_J(x) := (I - \mathrm{D}\boldsymbol{g}^{\mathcal{T}}(\mathrm{D}\boldsymbol{g}\mathrm{D}\boldsymbol{g}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{g})\nabla J(x),$$
$$\boldsymbol{\xi}_C(x) := \mathrm{D}\boldsymbol{g}^{\mathcal{T}}(\mathrm{D}\boldsymbol{g}\mathrm{D}\boldsymbol{g}^{\mathcal{T}})^{-1}\boldsymbol{g}(x).$$

## Proposition 1

*Assume that the constraints* $\boldsymbol{g}$ *are qualified and consider the flow*

$$\begin{cases} \dot{x} = -\alpha_J(I - \mathrm{D}\boldsymbol{g}^{\mathcal{T}}(\mathrm{D}\boldsymbol{g}\mathrm{D}\boldsymbol{g}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{g}(x))\nabla J(x) - \alpha_C \mathrm{D}\boldsymbol{g}^{\mathcal{T}}(\mathrm{D}\boldsymbol{g}\mathrm{D}\boldsymbol{g}^{\mathcal{T}})^{-1}\boldsymbol{g}(x) \\ x(0) = x_0 \end{cases} \quad (2)$$

*for some* $\alpha_J, \alpha_C > 0$*. Then the following properties hold true:*

# Equality constrained optimization

## Proposition 1

*Assume that the constraints* $\mathbf{g}$ *are qualified and consider the flow*

$$\begin{cases} \dot{x} = -\alpha_J(I - \mathrm{D}\mathbf{g}^{\mathcal{T}}(\mathrm{D}\mathbf{g}\mathrm{D}\mathbf{g}^{\mathcal{T}})^{-1}\mathrm{D}\mathbf{g}(x))\nabla J(x) - \alpha_C \mathrm{D}\mathbf{g}^{\mathcal{T}}(\mathrm{D}\mathbf{g}\mathrm{D}\mathbf{g}^{\mathcal{T}})^{-1}\mathbf{g}(x) \\ x(0) = x_0 \end{cases} \tag{2}$$

*for some* $\alpha_J, \alpha_C > 0$. *Then the following properties hold true:*

1. *The violation of the constraints decreases exponentially:*

$$\forall t \in [0, T],\ \mathbf{g}(x(t)) = e^{-\alpha_C t}\mathbf{g}(x_0).$$

2. $J(x(t))$ *decreases "as soon as the violation of the constraints is sufficiently small":*

$$\forall t \in [0, T],\ ||\Pi_{g(x)}(\nabla J(x(t)))||_V^2 > Ce^{-\alpha_C t} \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}J(x(t)) < 0.$$

3. *Any stationary point* $x^*$ *of eq. (2) satisfies the first-order KKT conditions, that is:*

$$\begin{cases} \mathbf{g}(x^*) = 0 \\ \exists \boldsymbol{\lambda}^* \in \mathbb{R}^p,\ \nabla J(x^*) + \mathrm{D}\mathbf{g}^{\mathcal{T}}(x^*)\boldsymbol{\lambda}^* = \Pi_{g(x^*)}(\nabla J(x^*)) = 0. \end{cases}$$

# Equality constrained optimization

Today: we see how to solve equality *and inequality* constrained optimization problems:

$$\min_{x \in V} \quad J(x) \qquad \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \le i \le p \\ h_j(x) \le 0, & 1 \le j \le q, \end{cases} \tag{3}$$

# Equality constrained optimization

Today: we see how to solve equality *and inequality* constrained optimization problems:

$$\min_{x \in V} \quad J(x) \\ \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) \leq 0, & 1 \leq j \leq q, \end{cases} \tag{3}$$

- It is *possible* to recast eq. (3) as an equality constrained optimization problem using artificial **slack variables**:

$$\min_{x \in V, (z_j)_{1 \leq j \leq q}} \quad J(x) \\ \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) + z_j^2 = 0 & 1 \leq i \leq q, \end{cases} \tag{4}$$

# Equality constrained optimization

Today: we see how to solve equality *and inequality* constrained optimization problems:

$$\min_{x\in V} \quad J(x)$$
$$\text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) \leq 0, & 1 \leq j \leq q, \end{cases} \tag{3}$$

- It is *possible* to recast eq. (3) as an equality constrained optimization problem using artificial **slack variables**:

$$\min_{x\in V, (z_j)_{1\leq j\leq q}} \quad J(x)$$
$$\text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) + z_j^2 = 0 & 1 \leq i \leq q, \end{cases} \tag{4}$$

- However, it is also possible to solve eq. (3) directly.

# Equality constrained optimization

Today: we see how to solve equality *and inequality* constrained optimization problems:

$$\min_{x\in V} \quad J(x) \qquad \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \le i \le p \\ h_j(x) \le 0, & 1 \le j \le q, \end{cases} \tag{3}$$

- It is *possible* to recast eq. (3) as an equality constrained optimization problem using artificial **slack variables**:

$$\min_{x\in V, (z_j)_{1\le j\le q}} \quad J(x) \qquad \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \le i \le p \\ h_j(x) + z_j^2 = 0 & 1 \le i \le q, \end{cases} \tag{4}$$

- However, it is also possible to solve eq. (3) directly.

# Equality constrained optimization

Today: we see how to solve equality *and inequality* constrained optimization problems:

$$\min_{x\in V} \quad J(x) \qquad \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) \leq 0, & 1 \leq j \leq q, \end{cases} \tag{3}$$

- It is *possible* to recast eq. (3) as an equality constrained optimization problem using artificial **slack variables**:

$$\min_{x\in V, (z_j)_{1\leq j\leq q}} \quad J(x) \qquad \text{s.t.} \begin{cases} g_i(x) = 0, & 1 \leq i \leq p \\ h_j(x) + z_j^2 = 0 & 1 \leq i \leq q, \end{cases} \tag{4}$$

- However, it is also possible to solve eq. (3) directly.

Inequality constraints have a **fully different nature** than equality constraints.

## Outline


1. Reminders on smooth constrained optimization
2. Gradient flows for unconstrained optimization
3. Constrained optimization:
   3.1 Extension to equality constrained optimization
   3.2 Extension to equality and inequality constrained optimization
4. Numerical implementation
5. Numerical examples

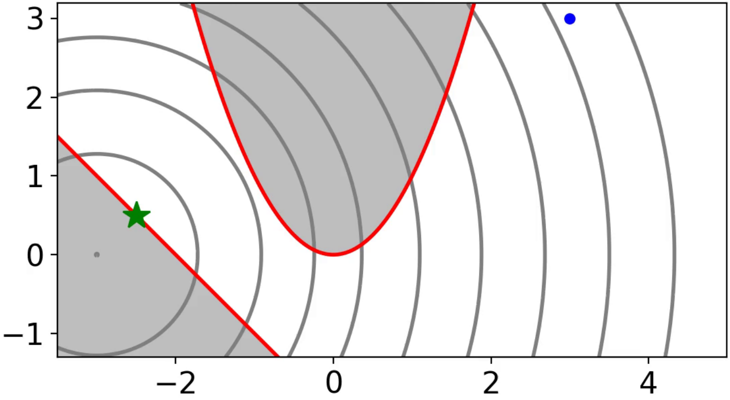
# Outline

1. Reminders on smooth constrained optimization
2. Gradient flows for unconstrained optimization
3. Constrained optimization:
    3.1 Extension to equality constrained optimization
    3.2 Extension to equality and inequality constrained optimization
4. Numerical implementation
5. Numerical examples

# Null space gradient flows for constrained optimization

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \quad J(x_1,x_2) = x_1^2 + (x_2+3)^2$$

$$\text{s.t.} \quad \begin{cases} h_1(x_1,x_2) = -x_1^2 + x_2 & \leq 0 \\ h_2(x_1,x_2) = -x_1 - x_2 - 2 & \leq 0 \end{cases}$$



# Equality and inequality constrained optimization

For *both* equality constraints $\boldsymbol{g}(x) = 0$ and inequality constraints $\boldsymbol{h}(x) \leq 0$, we consider:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x(t)) - \alpha_C \boldsymbol{\xi}_C(x(t))$$

with

$$\boldsymbol{\xi}_J(x) := (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}(x)}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}(x)})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\boldsymbol{\xi}_C(x) = \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

# Equality and inequality constrained optimization

For *both* equality constraints $\boldsymbol{g}(x) = 0$ and inequality constraints $\boldsymbol{h}(x) \leq 0$, we consider:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x(t)) - \alpha_C \boldsymbol{\xi}_C(x(t))$$

with

$$\boldsymbol{\xi}_J(x) := (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\boldsymbol{\xi}_C(x) = \mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

$\widetilde{I}(x)$ the set of violated constraints:

$$\widetilde{I}(x) = \{i \in \{1, \dots, q\} \mid h_i(x) \geqslant 0\}.$$

$$\boldsymbol{C}_{\widetilde{I}(x)} = \left[\boldsymbol{g}(x) \quad | \quad (h_i(x))_{i \in \widetilde{I}(x)}\right]^T$$

# Equality and inequality constrained optimization

For *both* equality constraints $\boldsymbol{g}(x) = 0$ and inequality constraints $\boldsymbol{h}(x) \leq 0$, we consider:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x(t)) - \alpha_C \boldsymbol{\xi}_C(x(t))$$

with

$$\boldsymbol{\xi}_J(x) := (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\boldsymbol{\xi}_C(x) = \mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

$\widehat{I}(x) \subset \widetilde{I}(x)$ is an "optimal" subset of the active or violated constraints which shall be computed by mean of a dual subproblem.

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

$$\boldsymbol{C}_{\widehat{I}(x)} = \left[\boldsymbol{g}(x) \quad | \quad (h_i(x))_{i \in \widehat{I}(x)}\right]^T$$

# Equality and inequality constrained optimization

For *both* equality constraints $\boldsymbol{g}(x) = 0$ and inequality constraints $\boldsymbol{h}(x) \leq 0$, we consider:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x(t)) - \alpha_C \boldsymbol{\xi}_C(x(t))$$

with

$$\boldsymbol{\xi}_J(x) := (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\boldsymbol{\xi}_C(x) = \mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

$\widehat{I}(x) \subset \widetilde{I}(x)$ is an "optimal" subset of the active or violated constraints which shall be computed by mean of a dual subproblem.

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

$$\boldsymbol{C}_{\widehat{I}(x)} = \left[\boldsymbol{g}(x) \quad | \quad (h_i(x))_{i \in \widehat{I}(x)}\right]^T$$

We assume the constraints to be qualified:

$$\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}} \text{ is invertible.}$$

# Equality and inequality constrained optimization

Consider the optimization problem

$$\min_{x \in V} \quad J(x)$$
$$\text{s.t.} \begin{cases} \boldsymbol{g}(x) = 0 \\ \boldsymbol{h}(x) \leq 0, \end{cases}$$

## Definition 2 (range space step)

The range step $\boldsymbol{\xi}_C(x)$ is defined by

$$\boldsymbol{\xi}_C(x) := \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

# Equality and inequality constrained optimization

Consider the optimization problem

$$\min_{x \in V} \quad J(x)$$
$$\text{s.t.} \begin{cases} \boldsymbol{g}(x) = 0 \\ \boldsymbol{h}(x) \leq 0, \end{cases}$$

## Definition 2 (range space step)

The range step $\boldsymbol{\xi}_C(x)$ is defined by

$$\boldsymbol{\xi}_C(x) := \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widetilde{I}(x)})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

In particular:

1. $\boldsymbol{\xi}_C(x)$ is orthogonal to $\mathrm{Ker}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)})$.

Consider the optimization problem

$$\min_{x\in V} \quad J(x)$$
$$\text{s.t.} \begin{cases} \boldsymbol{g}(x) = 0 \\ \boldsymbol{h}(x) \leq 0, \end{cases}$$

## Definition 2 (range space step)

The range step $\boldsymbol{\xi}_C(x)$ is defined by

$$\boldsymbol{\xi}_C(x) := \mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{\widetilde{I}(x)}(x),$$

In particular:

1. $\boldsymbol{\xi}_C(x)$ is orthogonal to $\mathrm{Ker}(\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)})$.
2. $-\boldsymbol{\xi}_C(x)$ is a Gauss-Newton direction for the violation of the constraints:

$$\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x)}(-\boldsymbol{\xi}_C(x)) = -\boldsymbol{C}_{\widetilde{I}(x)}(x).$$

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi} \in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \quad \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ ||\boldsymbol{\xi}||_V \leq 1. \end{cases} \tag{5}$$

# Equality and inequality constrained optimization

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi}\in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ ||\boldsymbol{\xi}||_V \leq 1. \end{cases} \tag{5}$$

$\boldsymbol{\xi}$ is tangent to the admissible cone tangent to the constraints

# Equality and inequality constrained optimization

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi}\in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \quad \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ ||\boldsymbol{\xi}||_V \leq 1. \end{cases} \tag{5}$$

$\boldsymbol{\xi}$ is tangent to the admissible cone tangent to the constraints

$\boldsymbol{\xi}^*(x)$ is the best "descent direction" respecting locally the constraints.

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi}\in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \quad \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ ||\boldsymbol{\xi}||_V \leq 1. \end{cases} \tag{5}$$

$\boldsymbol{\xi}$ is tangent to the admissible cone tangent to the constraints

$\boldsymbol{\xi}^*(x)$ is the best "descent direction" respecting locally the constraints.

In what follows, we give a characterization of $\boldsymbol{\xi}^*(x)$.

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi}\in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ \|\boldsymbol{\xi}\|_V \leq 1. \end{cases} \tag{5}$$

$\boldsymbol{\xi}$ is tangent to the admissible cone tangent to the constraints

$\boldsymbol{\xi}^*(x)$ is the best "descent direction" respecting locally the constraints.

In what follows, we give a characterization of $\boldsymbol{\xi}^*(x)$.

We call eq. (8) the **primal problem**.

## Proposition 2

*Let the constraints be satisfied. There exists a unique couple of multipliers* $\boldsymbol{\lambda}^*(x) \in \mathbb{R}^p$ *and* $\boldsymbol{\mu}^*(x) \in \mathbb{R}_+^{\widetilde{q}(x)}$ *solution to the following quadratic optimization problem which is the* dual *of eq.* (8) *:*

$$(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}||_V.$$

### Proposition 3

*Let*

$$m^*(x) := ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}^*(x)||_V$$

*be the value of the dual problem . Then the value of the primal problem is $p^*(x) = -m^*(x)$ and the following alternative holds:*

## Proposition 3

*Let*

$$m^*(x) := ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}}\boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}}\boldsymbol{\mu}^*(x)||_V$$

*be the value of the dual problem . Then the value of the primal problem is* $p^*(x) = -m^*(x)$ *and the following alternative holds:*

1. $m^*(x) = 0$: *the KKT conditions hold with (necessarily unique) Lagrange multipliers* $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}_+^{\widetilde{q}(x)}$:

$$\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}}\boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}}\boldsymbol{\mu}^*(x) = 0. \tag{6}$$

   *A minimizer of the primal problem is* $\boldsymbol{\xi}^*(x) = 0$.

## Proposition 3

*Let*

$$m^*(x) := ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}^*(x)||_V$$

*be the value of the dual problem . Then the value of the primal problem is* $p^*(x) = -m^*(x)$ *and the following alternative holds:*

1. $m^*(x) = 0$*: the KKT conditions hold with (necessarily unique) Lagrange multipliers* $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}_+^{\widetilde{q}(x)}$*:*

$$\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}^*(x) = 0. \tag{6}$$

   *A minimizer of the primal problem is* $\boldsymbol{\xi}^*(x) = 0$*.*

2. $m^*(x) > 0$ *: eq. (6) does not hold and there exists a unique minimizer* $\boldsymbol{\xi}^*(x)$ *to the primal problem, given by*

$$\boldsymbol{\xi}^*(x) = -\frac{\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}^*(x)}{||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda}^*(x) + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}^*(x)||_V}. \tag{7}$$

# Equality and inequality constrained optimization

Let $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}^{\widetilde{q}(x)}$ the solutions of the following dual minimization problem:

$$(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}||_V.$$

Let $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}^{\widetilde{q}(x)}$ the solutions of the following dual minimization problem:

$$(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}} \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}} \boldsymbol{\mu}||_V.$$

## Proposition 4

*Define $\widehat{I}(x)$ the set obtained by collecting the non zero components of $\boldsymbol{\mu}^*(x)$:*

$$\widehat{I}(x) := \{ i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0 \}.$$

Let $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}^{\widetilde{q}(x)}$ the solutions of the following dual minimization problem:

$$(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) := \arg \min_{\substack{\boldsymbol{\lambda}\in\mathbb{R}^p \\ \boldsymbol{\mu}\in\mathbb{R}^{\widetilde{q}(x)},\, \boldsymbol{\mu}\geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}}\boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}}\boldsymbol{\mu}||_V.$$

## Proposition 4

*Define $\widehat{I}(x)$ the set obtained by collecting the non zero components of $\boldsymbol{\mu}^*(x)$:*

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

*Then $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x))$ and $\boldsymbol{\xi}^*(x)$ are explicitly given in terms of $\widehat{I}(x)$ by:*

$$\begin{bmatrix} \boldsymbol{\lambda}^*(x) \\ \widehat{\boldsymbol{\mu}}^*(x) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\lambda}_{\widehat{I}(x)}(x) \\ \boldsymbol{\mu}_{\widehat{I}(x)}(x) \end{bmatrix} = -(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\nabla J(x),$$

$$\boldsymbol{\xi}^*(x) = -\frac{\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))}{||\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))||_V},$$

Let $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) \in \mathbb{R}^p \times \mathbb{R}^{\widetilde{q}(x)}$ the solutions of the following dual minimization problem:

$$(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x)) := \arg \min_{\substack{\boldsymbol{\lambda}\in\mathbb{R}^p \\ \boldsymbol{\mu}\in\mathbb{R}^{\widetilde{q}(x)},\ \boldsymbol{\mu}\geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^{\mathcal{T}}\boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)^{\mathcal{T}}\boldsymbol{\mu}||_V.$$

## Proposition 4

*Define $\widehat{I}(x)$ the set obtained by collecting the non zero components of $\boldsymbol{\mu}^*(x)$:*

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

*Then $(\boldsymbol{\lambda}^*(x), \boldsymbol{\mu}^*(x))$ and $\boldsymbol{\xi}^*(x)$ are explicitly given in terms of $\widehat{I}(x)$ by:*

$$\begin{bmatrix} \boldsymbol{\lambda}^*(x) \\ \widehat{\boldsymbol{\mu}}^*(x) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\lambda}_{\widehat{I}(x)}(x) \\ \boldsymbol{\mu}_{\widehat{I}(x)}(x) \end{bmatrix} = -(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\nabla J(x),$$

$$\boldsymbol{\xi}^*(x) = -\frac{\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))}{||\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))||_V},$$

*where $\widehat{\boldsymbol{\mu}}^*(x) := (\mu_i^*(x))_{i\in\widehat{I}(x)}$ is the vector collecting all positive components of $\boldsymbol{\mu}^*(x)$.*

The **null space step**, $-\boldsymbol{\xi}_J(x)$ shall be set positively proportional to the solution $\boldsymbol{\xi}^*(x)$ of the following minimization problem:

$$\boldsymbol{\xi}^*(x) = \arg\min_{\boldsymbol{\xi}\in V} \quad \mathrm{D}J(x)\boldsymbol{\xi}$$
$$\text{s.t.} \begin{cases} \mathrm{D}\boldsymbol{g}(x)\boldsymbol{\xi} = 0 \\ \mathrm{D}\boldsymbol{h}_{\widetilde{I}(x)}(x)\boldsymbol{\xi} \leq 0 \\ ||\boldsymbol{\xi}||_V \leq 1. \end{cases} \tag{8}$$

where $\boldsymbol{h}_{\widetilde{I}(x)}(x) = (h_i(x))_{i\in\widetilde{I}(x)}$.

# Equality and inequality constrained optimization

In other words, $\boldsymbol{\xi}^*(x)$ is explicitly given by:

$$\boldsymbol{\xi}^*(x) = -\frac{\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))}{||\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))||_V},$$

with

$$\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x)) = (I - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

In other words, $\boldsymbol{\xi}^*(x)$ is explicitly given by:

$$\boldsymbol{\xi}^*(x) = -\frac{\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))}{||\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))||_V},$$

with

$$\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x)) = (I - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

whence our definition of the null space step:

In other words, $\boldsymbol{\xi}^*(x)$ is explicitly given by:

$$\boldsymbol{\xi}^*(x) = -\frac{\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))}{||\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x))||_V},$$

with

$$\Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x)) = (I - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})(\nabla J(x))$$

$$\widehat{I}(x) := \{i \in \widetilde{I}(x) \,|\, \mu_i^*(x) > 0\}.$$

whence our definition of the null space step:

## Definition 3

The null space direction $\boldsymbol{\xi}_J(x)$ is defined by:

$$\boldsymbol{\xi}_J(x) := \Pi_{\boldsymbol{C}_{\widehat{I}(x)}}(\nabla J(x)) = (I - \mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}(x)^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}(x)})\nabla J(x),$$

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

We can prove similarly:

1. Constraints are asymptotically satisfied:

$$\boldsymbol{g}(x(t)) = e^{-\alpha_C t}\boldsymbol{g}(x(0)) \text{ and } \boldsymbol{h}_{\widetilde{I}(x(t))} \leq e^{-\alpha_C t}\boldsymbol{h}(x(0))$$

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

We can prove similarly:

1. Constraints are asymptotically satisfied:

$$\boldsymbol{g}(x(t)) = e^{-\alpha_C t}\boldsymbol{g}(x(0)) \text{ and } \boldsymbol{h}_{\widetilde{I}(x(t))} \leq e^{-\alpha_C t}\boldsymbol{h}(x(0))$$

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

We can prove similarly:

1. Constraints are asymptotically satisfied:

$$\boldsymbol{g}(x(t)) = e^{-\alpha_C t}\boldsymbol{g}(x(0)) \text{ and } \boldsymbol{h}_{\widetilde{I}(x(t))} \leq e^{-\alpha_C t}\boldsymbol{h}(x(0))$$

2. $J$ decreases as soon as the violation $\boldsymbol{C}_{\widetilde{I}(x(t))}$ is sufficiently small

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

We can prove similarly:

1. Constraints are asymptotically satisfied:

$$\boldsymbol{g}(x(t)) = e^{-\alpha_C t} \boldsymbol{g}(x(0)) \text{ and } \boldsymbol{h}_{\widetilde{I}(x(t))} \leq e^{-\alpha_C t} \boldsymbol{h}(x(0))$$

2. $J$ decreases as soon as the violation $\boldsymbol{C}_{\widetilde{I}(x(t))}$ is sufficiently small

# Equality and inequality constrained optimization

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

We can prove similarly:

1. Constraints are asymptotically satisfied:

$$\boldsymbol{g}(x(t)) = e^{-\alpha_C t}\boldsymbol{g}(x(0)) \text{ and } \boldsymbol{h}_{\widetilde{I}(x(t))} \leq e^{-\alpha_C t}\boldsymbol{h}(x(0))$$

2. $J$ decreases as soon as the violation $\boldsymbol{C}_{\widetilde{I}(x(t))}$ is sufficiently small
3. All stationary points $x^*$ of the ODE are KKT points

# Outline

1. Reminders on smooth constrained optimization
2. Gradient flows for unconstrained optimization
3. Constrained optimization:
   3.1 Extension to equality constrained optimization
   3.2 Extension to equality and inequality constrained optimization
4. Numerical implementation
5. Numerical examples

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

Solutions considered:

- Detect the constraints slightly before they get violated

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

Solutions considered:

- Detect the constraints slightly before they get violated
- Use an appropriate "merit function"

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

Solutions considered:

- Detect the constraints slightly before they get violated
- Use an appropriate "merit function"

# Numerical implementation issues

Consider the null space gradient flow:

$$\dot{x} = -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x).$$

- The right-hand side of the null space ODE is **discontinuous** when the set $\widehat{I}(x)$ changes.
- How to select a "sufficiently small" time step $\Delta t$ ?

Solutions considered:

- Detect the constraints slightly before they get violated
- Use an appropriate "merit function"

We discretize the flow with an explicit Euler scheme:

$$x_{n+1} = x_n - \Delta t_n(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)),$$

with $\Delta t_n$ an adaptive time-step.

Replace $\widetilde{I}(x_n)$ with $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

Replace $\widetilde{I}(x_n)$ with $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

Using a time step $\Delta t$, it is natural to set

$$\epsilon_i = ||\nabla h_i(x_n)||\Delta t$$

# Numerical implementation issues

## Feeling inequality constraints from a short distance

Replace $\widetilde{I}(x_n)$ with $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \dots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

Using a time step $\Delta t$, it is natural to set

$$\epsilon_i = ||\nabla h_i(x_n)||\Delta t$$

- compute the set of indices $\widehat{I}_\epsilon(x_n)$ of constraints "not aligned with the gradients" when crossing $\{\boldsymbol{h} = 0\}$

Replace $\widetilde{I}(x_n)$ with $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \dots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

Using a time step $\Delta t$, it is natural to set

$$\epsilon_i = ||\nabla h_i(x_n)||\Delta t$$

- compute the set of indices $\widehat{I}_\epsilon(x_n)$ of constraints "not aligned with the gradients" when crossing $\{\boldsymbol{h} = 0\}$
- compute $\boldsymbol{\xi}_{J,\epsilon}(x_n)$ and $\boldsymbol{\xi}_{C,\epsilon}(x_n)$ as follows:

$$\boldsymbol{\xi}_{J,\epsilon}(x_n) := (I - \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$

$$\boldsymbol{\xi}_{C,\epsilon}(x_n) := \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I^*_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{I^*_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I^*_\epsilon(x_n)})^{-1}\boldsymbol{C}_{I^*_\epsilon(x_n)}(x_n),$$

where $I^*_\epsilon(x_n) = \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$ is the set of constraints that are either violated, saturated or not aligned with the gradient.

Replace $\widetilde{I}(x_n)$ with $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \dots, q\} \,|\, h_i(x_n) \geqslant -\epsilon_i\}.$$

Using a time step $\Delta t$, it is natural to set

$$\epsilon_i = ||\nabla h_i(x_n)||\Delta t$$

- compute the set of indices $\widehat{I}_\epsilon(x_n)$ of constraints "not aligned with the gradients" when crossing $\{\boldsymbol{h} = 0\}$
- compute $\boldsymbol{\xi}_{J,\epsilon}(x_n)$ and $\boldsymbol{\xi}_{C,\epsilon}(x_n)$ as follows:

$$\boldsymbol{\xi}_{J,\epsilon}(x_n) := (I - \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$

$$\boldsymbol{\xi}_{C,\epsilon}(x_n) := \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I^*_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{I^*_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I^*_\epsilon(x_n)})^{-1}\boldsymbol{C}_{I^*_\epsilon(x_n)}(x_n),$$

  where $I^*_\epsilon(x_n) = \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$ is the set of constraints that are either violated, saturated or not aligned with the gradient.
- Including constraints of $\widehat{I}_\epsilon(x_n)$ not in $\widetilde{I}(x_n)$ further stabilizes these closer to the zero barrier.

## Lemma 4

*For a given $x_n \in V$, let $\mathtt{merit}_{x_n} : V \to \mathbb{R}$ be the function defined by*

$$\mathtt{merit}_{x_n}(x) := \alpha_J \left( J(x) + \boldsymbol{\Lambda}(x_n)^T \boldsymbol{C}_{\widetilde{I}(x_n)}(x) \right) + \frac{\alpha_C}{2} \boldsymbol{C}_{\widetilde{I}(x_n)}(x)^T \boldsymbol{S}(x_n) \boldsymbol{C}_{\widetilde{I}(x_n)}(x)$$

## Lemma 4

*For a given $x_n \in V$, let $\texttt{merit}_{x_n} : V \to \mathbb{R}$ be the function defined by*

$$\texttt{merit}_{x_n}(x) := \alpha_J \left( J(x) + \boldsymbol{\Lambda}(x_n)^T \boldsymbol{C}_{\widetilde{I}(x_n)}(x) \right) + \frac{\alpha_C}{2} \boldsymbol{C}_{\widetilde{I}(x_n)}(x)^T \boldsymbol{S}(x_n) \boldsymbol{C}_{\widetilde{I}(x_n)}(x)$$

*where* $\boldsymbol{\Lambda}(x_n) = \begin{bmatrix} \boldsymbol{\lambda}^*(x_n)^T & \boldsymbol{\mu}^*(x_n)^T \end{bmatrix}^T$ *is the vector of multipliers defined as the solution to the dual problem and* $\boldsymbol{S}(x_n) = (\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)^{\mathcal{T}})^{-1}$ *is* symmetric positive definite. *Then*

$$x_{n+1} = x_n - \Delta t_n(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)),$$

*is a gradient step for decreasing the function* $\texttt{merit}_{x_n}$,

## Lemma 4

*For a given* $x_n \in V$*, let* $\mathtt{merit}_{x_n} : V \to \mathbb{R}$ *be the function defined by*

$$\mathtt{merit}_{x_n}(x) := \alpha_J \left( J(x) + \mathbf{\Lambda}(x_n)^T \boldsymbol{C}_{\widetilde{I}(x_n)}(x) \right) + \frac{\alpha_C}{2} \boldsymbol{C}_{\widetilde{I}(x_n)}(x)^T \boldsymbol{S}(x_n) \boldsymbol{C}_{\widetilde{I}(x_n)}(x)$$

*where* $\mathbf{\Lambda}(x_n) = \left[ \boldsymbol{\lambda}^*(x_n)^T \quad \boldsymbol{\mu}^*(x_n)^T \right]^T$ *is the vector of multipliers defined as the solution to the dual problem and* $\boldsymbol{S}(x_n) = (\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)^{\mathcal{T}})^{-1}$ *is* symmetric positive definite*. Then*

$$x_{n+1} = x_n - \Delta t_n(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)),$$

*is a gradient step for decreasing the function* $\mathtt{merit}_{x_n}$*, namely:*

$$\nabla \mathtt{merit}_{x_n}(x_n) = \alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n).$$

## Lemma 4

*For a given $x_n \in V$, let $\texttt{merit}_{x_n} : V \to \mathbb{R}$ be the function defined by*

$$\texttt{merit}_{x_n}(x) := \alpha_J \left( J(x) + \boldsymbol{\Lambda}(x_n)^T \boldsymbol{C}_{\widetilde{I}(x_n)}(x) \right) + \frac{\alpha_C}{2} \boldsymbol{C}_{\widetilde{I}(x_n)}(x)^T \boldsymbol{S}(x_n) \boldsymbol{C}_{\widetilde{I}(x_n)}(x)$$

*where $\boldsymbol{\Lambda}(x_n) = \left[ \boldsymbol{\lambda}^*(x_n)^T \quad \boldsymbol{\mu}^*(x_n)^T \right]^T$ is the vector of multipliers defined as the solution to the dual problem and $\boldsymbol{S}(x_n) = (\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n) \mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)^{\mathcal{T}})^{-1}$ is* symmetric positive definite*. Then*

$$x_{n+1} = x_n - \Delta t_n (\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)),$$

*is a gradient step for decreasing the function $\texttt{merit}_{x_n}$, namely:*

$$\nabla \texttt{merit}_{x_n}(x_n) = \alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n).$$

Consequently:

- if $\texttt{merit}_{x_n}(x_{n+1}) > \texttt{merit}_{x_n}(x_n)$, then $\Delta t_n$ is too large !

## Lemma 4

*For a given $x_n \in V$, let $\mathtt{merit}_{x_n} : V \to \mathbb{R}$ be the function defined by*

$$\mathtt{merit}_{x_n}(x) := \alpha_J \left( J(x) + \boldsymbol{\Lambda}(x_n)^T \boldsymbol{C}_{\widetilde{I}(x_n)}(x) \right) + \frac{\alpha_C}{2} \boldsymbol{C}_{\widetilde{I}(x_n)}(x)^T \boldsymbol{S}(x_n) \boldsymbol{C}_{\widetilde{I}(x_n)}(x)$$

*where $\boldsymbol{\Lambda}(x_n) = \begin{bmatrix} \boldsymbol{\lambda}^*(x_n)^T & \boldsymbol{\mu}^*(x_n)^T \end{bmatrix}^T$ is the vector of multipliers defined as the solution to the dual problem and $\boldsymbol{S}(x_n) = (\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)\mathrm{D}\boldsymbol{C}_{\widetilde{I}(x_n)}(x_n)^{\mathcal{T}})^{-1}$ is* symmetric positive definite*. Then*

$$x_{n+1} = x_n - \Delta t_n(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)),$$

*is a gradient step for decreasing the function $\mathtt{merit}_{x_n}$, namely:*

$$\nabla \mathtt{merit}_{x_n}(x_n) = \alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n).$$

Consequently:

- if $\mathtt{merit}_{x_n}(x_{n+1}) > \mathtt{merit}_{x_n}(x_n)$, then $\Delta t_n$ is too large !
- in practice, one decreases $\Delta t$ a finite number of times until $\mathtt{merit}_{x_n}(x_{n+1}) < \mathtt{merit}_{x_n}(x_n)$.

# Summary

For $n = 1 \dots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.

# Summary

For $n = 1 \dots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.
2. For all inequality constraints $1 \leq i \leq q$, compute the tolerance

$$\epsilon_i := ||\nabla h_i(x_n)||_V \, \texttt{h}.$$

# Summary

For $n = 1 \ldots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.
2. For all inequality constraints $1 \leq i \leq q$, compute the tolerance

$$\epsilon_i := ||\nabla h_i(x_n)||_V \, \mathtt{h}.$$

3. Determine the set $\widetilde{I}(x_n)$ of active or violated constraints and the set $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant 0\}$$

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

# Summary

For $n = 1 \dots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.
2. For all inequality constraints $1 \leq i \leq q$, compute the tolerance
$$\epsilon_i := ||\nabla h_i(x_n)||_V \, \mathtt{h}.$$
3. Determine the set $\widetilde{I}(x_n)$ of active or violated constraints and the set $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":
$$\widetilde{I}(x_n) = \{i \in \{1, \dots, q\} \mid h_i(x_n) \geqslant 0\}$$
$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \dots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$
4. Denote by $\widetilde{q}_\epsilon := \#(\widetilde{I}_\epsilon)$. Solve the dual problem
$$(\boldsymbol{\lambda}_\epsilon^*(x_n), \boldsymbol{\mu}_\epsilon^*(x_n)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}_\epsilon(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^T \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}_\epsilon(x)}(x)^T \boldsymbol{\mu}||_V$$
to obtain the optimal Lagrange multiplier $\boldsymbol{\mu}^*(x_n)$.

# Summary

For $n = 1 \ldots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.
2. For all inequality constraints $1 \leq i \leq q$, compute the tolerance

$$\epsilon_i := ||\nabla h_i(x_n)||_V \, \mathtt{h}.$$

3. Determine the set $\widetilde{I}(x_n)$ of active or violated constraints and the set $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":

$$\widetilde{I}(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant 0\}$$

$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \ldots, q\} \mid h_i(x_n) \geqslant -\epsilon_i\}.$$

4. Denote by $\widetilde{q}_\epsilon := \#(\widetilde{I}_\epsilon)$. Solve the dual problem

$$(\boldsymbol{\lambda}_\epsilon^*(x_n), \boldsymbol{\mu}_\epsilon^*(x_n)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}_\epsilon(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^T \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}_\epsilon(x)}(x)^T \boldsymbol{\mu}||_V$$

to obtain the optimal Lagrange multiplier $\boldsymbol{\mu}^*(x_n)$.

## Summary

For $n = 1 \ldots \texttt{maxiter}$:

1. Compute the gradients $\nabla J(x_n)$, $\nabla g_i(x_n)$ and $\nabla h_j(x_n)$ for $1 \leq i \leq p$, $1 \leq j \leq q$ by solving, if necessary, the identification problem $a(\nabla J(x_n), \boldsymbol{\xi}) = \mathrm{D}J(x_n) \cdot \boldsymbol{\xi}$.
2. For all inequality constraints $1 \leq i \leq q$, compute the tolerance
$$\epsilon_i := ||\nabla h_i(x_n)||_V \, \texttt{h}.$$
3. Determine the set $\widetilde{I}(x_n)$ of active or violated constraints and the set $\widetilde{I}_\epsilon(x_n)$ of constraints violated "up to $\epsilon_i$":
$$\widetilde{I}(x_n) = \{i \in \{1, \ldots, q\} \,|\, h_i(x_n) \geqslant 0\}$$
$$\widetilde{I}_\epsilon(x_n) = \{i \in \{1, \ldots, q\} \,|\, h_i(x_n) \geqslant -\epsilon_i\}.$$
4. Denote by $\widetilde{q}_\epsilon := \#(\widetilde{I}_\epsilon)$. Solve the dual problem
$$(\boldsymbol{\lambda}_\epsilon^*(x_n), \boldsymbol{\mu}_\epsilon^*(x_n)) := \arg \min_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^p \\ \boldsymbol{\mu} \in \mathbb{R}^{\widetilde{q}_\epsilon(x)},\, \boldsymbol{\mu} \geqslant 0}} ||\nabla J(x) + \mathrm{D}\boldsymbol{g}(x)^T \boldsymbol{\lambda} + \mathrm{D}\boldsymbol{h}_{\widetilde{I}_\epsilon(x)}(x)^T \boldsymbol{\mu}||_V$$
to obtain the optimal Lagrange multiplier $\boldsymbol{\mu}^*(x_n)$.Infer the subset $\widehat{I}_\epsilon(x_n) \subset \widetilde{I}_\epsilon(x_n)$ indicating which constraints must remain active:
$$\widehat{I}_\epsilon(x_n) = \{i \in \widetilde{I}_\epsilon(x_n) \,|\, \mu_{\epsilon,i}^*(x_n) > \texttt{tolLag}\}. \tag{9}$$

# Summary

For $n = 1 \ldots \texttt{maxiter}$:

5. Let $I_\epsilon^*(x_n) := \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$. Form the constraint vectors $\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}(x_n)$ and $\boldsymbol{C}_{I_\epsilon^*(x_n)}(x_n)$ and compute

$$\boldsymbol{\xi}_J(x_n) = (I - \mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$

$$\boldsymbol{\xi}_C(x_n) = \mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{I_\epsilon^*(x_n)}.$$

# Summary

For $n = 1 \ldots \texttt{maxiter}$:

5. Let $I_\epsilon^*(x_n) := \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$. Form the constraint vectors $\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}(x_n)$ and $\boldsymbol{C}_{I_\epsilon^*(x_n)}(x_n)$ and compute
$$\boldsymbol{\xi}_J(x_n) = (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}^{\mathcal{T}})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$
$$\boldsymbol{\xi}_C(x_n) = \mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}^{\mathcal{T}}(\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}^{\mathcal{T}})^{-1}\boldsymbol{C}_{I_\epsilon^*(x_n)}.$$
6. For $k = 1 \ldots \texttt{maxtrials}$,

# Summary

For $n = 1 \ldots \texttt{maxiter}$:

5. Let $I_\epsilon^*(x_n) := \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$. Form the constraint vectors $\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}(x_n)$ and $\boldsymbol{C}_{I_\epsilon^*(x_n)}(x_n)$ and compute

$$\boldsymbol{\xi}_J(x_n) = (\boldsymbol{I} - \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$

$$\boldsymbol{\xi}_C(x_n) = \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I_\epsilon^*(x_n)}(\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I_\epsilon^*(x_n)})^{-1}\boldsymbol{C}_{I_\epsilon^*(x_n)}.$$

6. For $k = 1 \ldots \texttt{maxtrials}$,
   6.1 Compute the step

$$x_{n+1} = x_n - \frac{\Delta t}{2^{k-1}}(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)).$$

# Summary

For $n = 1 \dots \texttt{maxiter}$:

5. Let $I_\epsilon^*(x_n) := \widetilde{I}(x_n) \cup \widehat{I}_\epsilon(x_n)$. Form the constraint vectors $\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}(x_n)$ and $\boldsymbol{C}_{I_\epsilon^*(x_n)}(x_n)$ and compute

$$\boldsymbol{\xi}_J(x_n) = (I - \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)}(\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{\widehat{I}_\epsilon(x_n)})^{-1}\mathrm{D}\boldsymbol{C}_{\widehat{I}_\epsilon(x_n)})\nabla J(x_n),$$

$$\boldsymbol{\xi}_C(x_n) = \mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I_\epsilon^*(x_n)}(\mathrm{D}\boldsymbol{C}_{I_\epsilon^*(x_n)}\mathrm{D}\boldsymbol{C}^{\mathcal{T}}_{I_\epsilon^*(x_n)})^{-1}\boldsymbol{C}_{I_\epsilon^*(x_n)}.$$

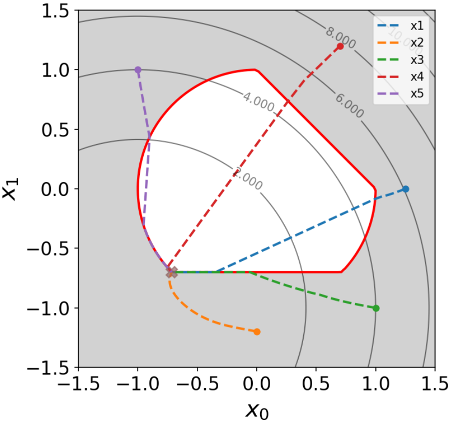
6. For $k = 1 \dots \texttt{maxtrials}$,
   6.1 Compute the step
   $$x_{n+1} = x_n - \frac{\Delta t}{2^{k-1}}(\alpha_J \boldsymbol{\xi}_J(x_n) + \alpha_C \boldsymbol{\xi}_C(x_n)).$$
   6.2 If $\texttt{merit}_{x_n}(x_{n+1}) < \texttt{merit}_{x_n}(x_n)$, then **break**

# Outline

1. Reminders on smooth constrained optimization
2. Gradient flows for unconstrained optimization
3. Constrained optimization:
   3.1 Extension to equality constrained optimization
   3.2 Extension to equality and inequality constrained optimization
4. Numerical implementation
5. Numerical examples

# Open source package available

Try it yourself!




`https://gitlab.com/florian.feppon/null-space-optimizer`

`pip install nullspace_optimizer`

# Basic problem 1

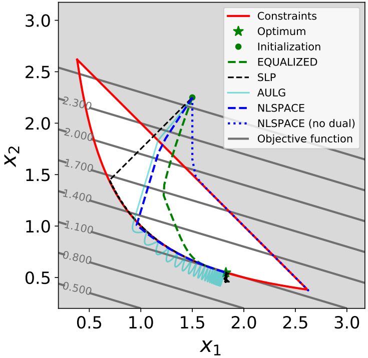
$$\min_{(x_1,x_2)\in\mathbb{R}^2} \quad J(x_1,x_2) := x_2 + 0.3x_1$$

$$\text{s.t.} \quad \begin{cases} h_1(x_1,x_2) := -x_2 + \dfrac{1}{x_1} & \leq 0, \\ h_2(x_1,x_2) := x_1 + x_2 - 3 & \leq 0. \end{cases}$$

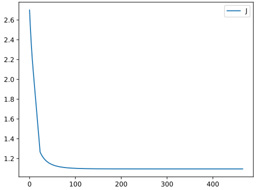
# Basic problem 1

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \quad J(x_1,x_2) := x_2 + 0.3x_1$$

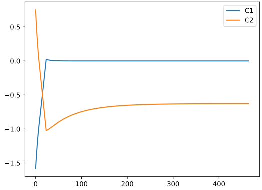
$$\text{s.t.} \quad \begin{cases} h_1(x_1,x_2) := -x_2 + \dfrac{1}{x_1} & \leq 0, \\ h_2(x_1,x_2) := x_1 + x_2 - 3 & \leq 0. \end{cases}$$



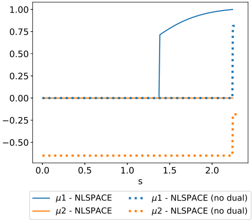
# Basic problem 1




(a) Objective function $J$




(b) Constraints $\boldsymbol{h}$




(c) Evolution of the Lagrange multipliers $\mu_1(x(s))$, $\mu_2(x(s))$

# Basic problem 1

More examples in python.