20.2 Constrained Optimization

① Definitions

- objective function
- minimization function
- performance index

• Formulation $\min_{x \in \mathbb{R}^n} f(x)$ subject to $\begin{cases} c_i(x) = 0 & i \in \mathcal{E} \\ c_i(x) \geq 0 & i \in \mathcal{I} \end{cases}$

- Equality constraints $c_i(x)$ s.t. $i \in \mathcal{E}$
- Inequality constraints $c_i(x)$ s.t. $i \in \mathcal{I}$
- Feasible set $\Omega = \{ x \text{ s.t. } c_i(x) = 0,\ i \in \mathcal{E};\ c_i(x) \geq 0,\ i \in \mathcal{I} \}$

compact formulation $\min_{x \in \Omega} f(x)$

• Optimality conditions (from Unconstrained Optimization)

① Necessary: must be satisfied by any solution point

② Sufficient: if satisfied at a point x^*, than x^* is a solution

① $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ positive semidefinite

② $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ positive definite

↳ is a strong local minimizer

• Local solutions (we need these definitions)

DEF: x^* is local solution if $x^* \in \Omega$ and $\exists$ a neighborhood $\mathcal{N}$ of x^* s.t. $f(x) \geq f(x^*)$, $\forall x \in \mathcal{N} \cap \Omega$

DEF: x^* is isolated local solution if $x^* \in \Omega$ and $\exists$ a neighborhood $\mathcal{N}$ of x^* s.t. x^* is the only local solution in $\mathcal{N} \cap \Omega$

DEF: x^* is a strong local solution if $x^* \in \Omega$ and $\exists$ a neighborhood $\mathcal{N}$ of x^* s.t.

$$f(x) > f(x^*) \quad \forall x \in \mathcal{N} \cap \Omega \text{ with } x \neq x^*$$

EX: How to define global solution?

↳ $x^* \in \Omega$ s.t. $f(x) \geq f(x^*) \ \forall x \in \Omega$

- Smoothness

↳ Important to allow algorithms to find the good search direction

↳ But! Non smooth unconstrained pb. $\overset{\text{sometimes}}{\Rightarrow}$ smooth constrained pb.

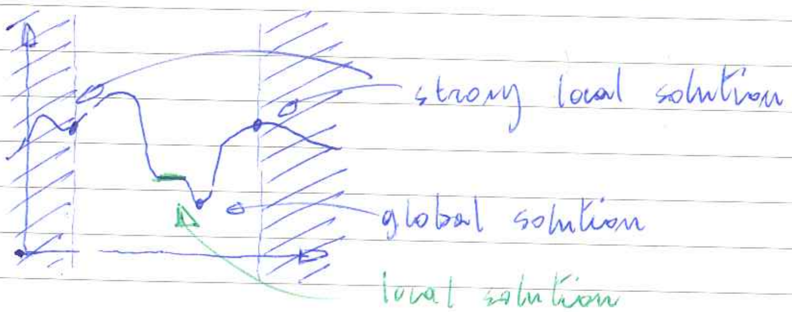
$$\min\{f_1(x), f_2(x)\} \geq a \iff \begin{cases} f_1(x) \geq a \\ f_2(x) \geq a \end{cases}$$

EX: $\min(x^2, x)$ $\Rightarrow$ non smooth! kink at $x = 0 = x^*$ but also the solution

$\Downarrow$

$$\min t \quad \text{s.t.} \begin{cases} t \geq x \\ t \geq x^2 \end{cases} \Rightarrow \text{smooth!}$$

EX:



Important

- Convex Hint: line between any two points lies above the graph, i.e., x^p s.t. $p \geq 1$ is convex

20.2.13

DEF: a pb. is convex if
- minimization function is convex
- feasible set Ω is convex
  - ↳ if all inequalities are concave

- constraints active/inactive (1)
  - ↳ active set $\mathcal{A}(x) = \mathcal{E} \cup \{ i \in \mathcal{I} \text{ s.t. } c_i(x) = 0 \}$

  $\mathcal{E}$: equality indices

  inequality indices for whose hold $c_i(x) = 0$, not only $c_i(x) \geq 0$

---

(2) 1st Order Optimality

- constraint qualification (2)

DEF: given feasible point x, $\mathcal{A}(x)$
set of linearized feasible directions is

$$\mathcal{F}(x) = \left\{ d \text{ s.t. } \begin{array}{ll} d^T \nabla c_i(x) = 0 & \forall i \in \mathcal{E} \\ d^T \nabla c_i(x) \geq 0 & \forall i \in \mathcal{A}(x) \cap \mathcal{I} \end{array} \right\}$$

Hint: a cone where is good to search for solution

⇓

In exercises often:
- ↳ LICQ Linear Independence constraint qualification
  holds if the set of active constraint gradients
  $\{ \nabla c_i(x) \text{ s.t. } i \in \mathcal{A}(x) \}$ is linearly indipendent
- ↳ linear constraints holds if
  $\{ \nabla c_i(x) \text{ s.t. } i \in \mathcal{A}(x) \}$ are linear function

- Lagrange function

DEF: $L(x,\lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$

Lagrange multiplier

- First-Order Necessary conditions / KKT Conditions

TH: suppose x^* is a local minimum

f, c_i are continuously differentiable

$\exists$ a Lagrange multiplier vector λ^* with components λ_i^*, $i \in \mathcal{E} \cup \mathcal{I}$ s.t.

stationarity	i. $\nabla_x L(x^*, \lambda^*) = 0$		Karush-Kuhn-Tucker conditions
primal feasibility	ii. $c_i(x^*) = 0$	$\forall i \in \mathcal{E}$	
	iii. $c_i(x^*) \geq 0$	$\forall i \in \mathcal{I}$	
dual feasibility	iv. $\lambda_i^* \geq 0$	$\forall i \in \mathcal{I}$	
complementarity	v. $\lambda_i^* c_i(x^*) = 0$	$\forall i \in \mathcal{E} \cup \mathcal{I}$	

We can write i

$$0 = \nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - \sum_{i \in A(x^*)} \lambda_i^* \nabla c_i(x^*)$$

DIM: since $c_i(x^*) > 0 \quad \forall i \notin A(x^*)$ to have

$$\lambda_i^* c_i(x^*) = 0 \Rightarrow \lambda_i^* = 0$$

Hint: the TH verifies a solution point x^* (λ^* has to be found), just find λ^* s.t. $0 = \nabla_x L(x^*, \lambda^*)$ is true!

$\Leftrightarrow$

KKT stationarity $\Leftrightarrow \nabla_x L(x^*, \lambda^*) = 0$

EX 1 $\min \; x_1 + 2x_2$ s.t. $2 - x_1^2 - x_2^2 \geq 0$, $x_2 \geq 0$

a. Find the optimal point

$$x^* = \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}$$

Hint: can be found watching the pb. statement

b. Check the KKT conditions at the optimal point

the Lagrangian from the pb. is

$$L(x,\lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x) = x_1 + 2x_2 - \lambda_1 c_1(x) - \lambda_2 c_2(x)$$

with: $c_1(x) = 2 - x_1^2 - x_2^2$
$c_2(x) = x_2$ $\mathcal{I} = \{1, 2\}$

$$\nabla L_x(x^*, \lambda^*) = \begin{pmatrix} 1 + 2\lambda_1^* x_1^* \\ 2 + 2\lambda_1^* x_2^* - \lambda_2^* \end{pmatrix} = 0 \Rightarrow \lambda^* = \begin{pmatrix} \frac{1}{2\sqrt{2}} \\ 2 \end{pmatrix}$$

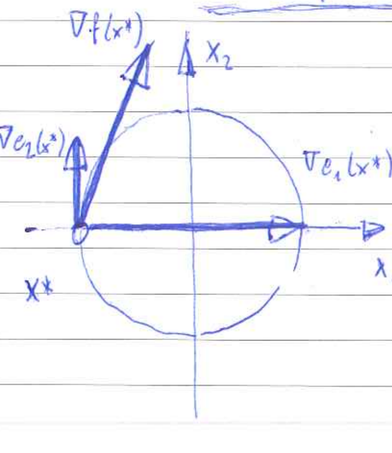
optimal point

$\hookrightarrow 1 + 2\lambda_1^*(-\sqrt{2}) = 0 \quad \lambda_1^* = \frac{1}{2\sqrt{2}}$

$\hookrightarrow 2 + 2\lambda_1^*(0) - \lambda_2^* = 0 \quad \lambda_2^* = 2$

With λ^*, all the KKT conditions are satisfied

c. Illustrate the gradients of the active constraints and the objective function at the optimal point(s)



$\nabla f(x) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \nabla f(x^*) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$\hookrightarrow \frac{\partial f(x)}{\partial x_1} = 1, \quad \frac{\partial f(x)}{\partial x_2} = 2$

$\nabla c_1(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} \quad \nabla c_1(x^*) = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$

$\hookrightarrow \frac{\partial c_1(x)}{\partial x_1} = -2x_1, \quad \frac{\partial c_1(x)}{\partial x_2} = -2x_2$

$\nabla c_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\hookrightarrow \frac{\partial c_2(x)}{\partial x_1} = 0 \quad \frac{\partial c_2(x)}{\partial x_2} = 1$

Hint: gradient of the objective for x_1 "goes up" by 2 (variation on x_2 axis) and "on right" by 1 (x_1 axis)

d. Explain why the Lagrange multiplier are positive

Hint: recall KKT conditions TH on p.4!

The pb. has 2 constraints $\{1,2\} \in \mathcal{I}$ (inequality)

$c_1(x) = 2 - x_1^2 - x_2^2$

$c_2(x) = x_2$

The (iv) of KKT: $\lambda_i^* \geq 0 \quad \forall i \in \mathcal{I}$, otherwise KKT doesn't hold

For this EX

(i) $\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) - \lambda_2^* \nabla c_2(x^*) = 0$

a weighted sum of 2 vectors with weight λ_1^*, λ_2^*

$\Downarrow$

$\nabla f(x^*) > 0 \iff \nabla c_1(x^*) > 0$ and $\nabla c_2(x^*) > 0$

$\Downarrow$

$\lambda_i > 0$ s.t. $i \in \mathcal{I} = \{1, 2\}$

Hint: a negative λ would point the gradient vector in opposite direction

Hint: all constraint gradients must have $\nabla f(x)$ angle less than 90°

e. Is this pb. a convex pb.? Hint: recall p.3!

$f(x_1, x_2) = x_1 + 2x_1$ is convex (all linear f. are)

$c_1(x) = 2 - x_1^2 - x_2^2 \geq 0$ is concave

$c_2(x) = x_2 \geq 0$ is concave

$\}$ Feasible set Ω is convex

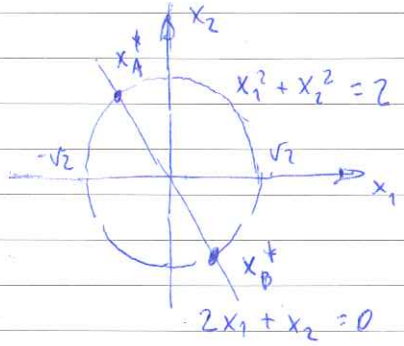
the pb. is convex

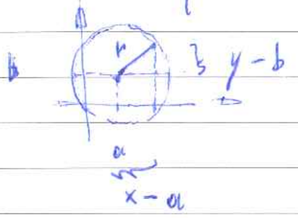
20.2.17

EX 2 $\min\ 2x_1 + x_2 \quad s.t.\ x_1^2 + x_2^2 - 2 = 0$

a. Find all extreme points

Hint: circle equation





$(x-a)^2 + (y-b)^2 = r^2$

to find the extreme points: $\begin{cases} x_1^2 + x_2^2 = 2 \\ 2x_1 + x_2 = 0 \end{cases}$ $\begin{cases} \dots \\ x_2 = -2x_1 \end{cases}$

$\begin{cases} x_1^2 + 4x_1^2 = 2 \\ \dots \end{cases}$ $\begin{cases} x_1 = \pm\sqrt{\frac{2}{5}} \\ x_2 = \mp 2\sqrt{\frac{2}{5}} \end{cases}$ $x_A^* = \left(-\sqrt{\frac{2}{5}}, 2\sqrt{\frac{2}{5}}\right)^T$

$x_B^* = \left(\sqrt{\frac{2}{5}}, -2\sqrt{\frac{2}{5}}\right)^T$

b. Check the KKT conditions at the extreme points

The Lagrangian for the pb. is: $L(x,\lambda) = 2x_1 + x_2 - \lambda_1 c_1(x)$

$c_1(x) = x_1^2 + x_2^2 - 2, \quad \mathcal{E} = \{1\}$

$L(x,\lambda) = 2x_1 + x_2 - \lambda_1 x_1^2 - \lambda_1 x_2^2 + 2\lambda_1$

$\hookrightarrow \frac{\partial L(x,\lambda)}{\partial x_1} = 2 - 2\lambda_1 x_1 \quad \frac{\partial L(x,\lambda)}{\partial x_2} = 1 - 2\lambda_1 x_2$

$$\nabla_x L(x^*, \lambda^*) = \begin{pmatrix} \frac{\partial L(x^*,\lambda^*)}{\partial x_1^*} \\ \frac{\partial L(x^*,\lambda^*)}{\partial x_2^*} \end{pmatrix} = \begin{pmatrix} 2 - 2\lambda_1 x_1^* \\ 1 - 2\lambda_1 x_2^* \end{pmatrix} = 0$$

For x_A^* $\quad \nabla_x L(x_A^*, \lambda^*) = \begin{pmatrix} 2 + 2\lambda_1\sqrt{\frac{2}{5}} \\ 1 - 4\lambda_1\sqrt{\frac{2}{5}} \end{pmatrix} = 0 \iff \begin{cases} 2 + 2\lambda_1\sqrt{\frac{2}{5}} = 0 \\ 1 - 4\lambda_1\sqrt{\frac{2}{5}} = 0 \end{cases} \begin{cases} \lambda_1 = -\sqrt{\frac{5}{2}} \\ \lambda_1 = \frac{1}{4}\sqrt{\frac{5}{2}} \end{cases}$

cannot be $\nabla L = 0$! KKT not satisfied !

c. Illustrate the gradients of the active constraint and the objective function at the optimal points

d. What is the value of the Lagrangian multiplier? Is this consistent with the KKT conditions?

e. Check the 2ND Order conditions for the extreme points.

---

3) 2ND Order Optimality

- 2ND Order Optimality conditions
  - ↳ determines if a direction w s.t. $w^T \nabla f(x^*) = 0$ will increse / decrease f
  - ↳ $w \rightarrow$ "critical cone"

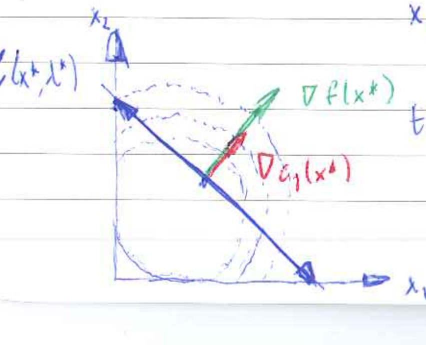
- Critical cone $\mathcal{C}$

linearized feasible directions

DEF $\mathcal{C}(x^*, \lambda^*) = \{ w \in \mathcal{F}(x^*) \text{ s.t. } \lambda_i^* \nabla c_i (x^*)^T w = 0$
$\forall i \in \mathcal{A}(x^*), \lambda_i^* > 0 \}$

Hint: $\lambda_i > 0 \Rightarrow$ the constrain $c_i(x)$ $i \in \mathcal{A}(x)$
($i \in \mathcal{A}(x) \Rightarrow$ will remain active)
even with small changes to J
($J \Rightarrow$ cost function)

EX: min $x_1^2 + x_2^2$ s.t.
$x_1 + x_2 - 3 \geq 0$



the critical cone is

$\mathcal{C}(x^*, \lambda^*) = h(-1, 1) \quad \forall h \in \mathbb{R}$

- **2ND Order conditions (Unconstrained pb.)**
  - TH: Necessary conditions for a week local minimum
    - i. $\nabla f(x^*) = 0$ is a stationary point
    - ii. $\nabla^2 f(x^*)$ is positive semi-definite
      - $w^T \nabla^2 f(x^*) w \geq 0 \quad \forall w \neq 0$
  - TH: Sufficient conditions for a strong local minimum
    - i. $\nabla f(x^*) = 0$ is a stationary point
    - ii. $\nabla^2 f(x^*) > 0$ is positive definite
      - $w^T \nabla^2 f(x^*) w > 0 \quad \forall w \neq 0$

- **2ND Order conditions (Constrained pb.)**
  - TH: Necessary condition for a week local minimum
    - i. KKT conditions holds
    - ii. $w^T \nabla^2 L(x^*, \lambda^*) w \geq 0 \quad \forall w \in \mathcal{G}(x^*, \lambda^*)$
  - TH: Sufficient condition for a strong local minimum
    - i. KKT conditions holds
    - ii. $w^T \nabla^2 L(x^*, \lambda^*) w > 0 \quad \forall w \in \mathcal{G}(x^*, \lambda^*)$

Hint: we know how to verify that x^* could be a solution (necessary condition) or is certainly a solution (sufficient condition)

- **Hessian**

$$H = \nabla^2 L = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}$$

20.2.10

EX 3 Problem 12.19 a, b and d in the book!

$$\min_{x \in \mathbb{R}^2} f(x) = -2x_1 + x_2 \quad \text{s.t.} \begin{cases} (1-x_1)^3 - x_2 \geq 0 \\ x_2 + 0.25\, x_1^2 - 1 \geq 0 \end{cases}$$

The optimal solution is $x^* = (0,1)^T \Rightarrow$ both $c_1(x)$ and $c_2(x)$ $\{1,2\} \in \mathcal{I}$ are active (equal to 0).

a. Do the LICQ holds for x^*? Hint: p. 3

if $\{\nabla c_i(x),\ i \in \mathcal{A}(x)\}$ is lineary indipendent

$$c_1(x) = (1-x_1)^3 - x_2 \qquad \nabla c_1(x) = \begin{pmatrix} -3(1-x)^2 \\ -1 \end{pmatrix} \qquad \nabla c_1(x^*) = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$$

$$\hookrightarrow \frac{\partial c_1(x)}{\partial x_1} = -3(+1-x)^2, \quad \frac{\partial c_1(x)}{\partial x_2} = -1$$

$$c_2(x) = x_2 + \frac{1}{4}x_1^2 - 1 \qquad \nabla c_2(x) = \begin{pmatrix} \frac{1}{2}x_1 \\ 1 \end{pmatrix} \qquad \nabla c_2(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\hookrightarrow \frac{\partial c_2(x)}{\partial x_1} = \frac{1}{2}x_1, \quad \frac{\partial c_2(x)}{\partial x_2} = 1$$

$$\begin{bmatrix} -3 & 0 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \text{Linear indipendence}$$

$$\hookrightarrow \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1} \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \checkmark$$

$\nabla c_1(x^*), \nabla c_2(x^*)$ are lin. indipendent LICQ holds

b. Are the KKT conditions satisfied?

The Lagrangian for the pb. is

$$L(x,\lambda) = -2x_1 + x_2 - \lambda_1 c_1(x) - \lambda_2 c_2(x) \quad \text{s.t.} \begin{cases} c_1(x) = (1-x_1)^3 - x_2 \\ c_2(x) = x_2 + 0.25x_1^2 - 1 \end{cases}$$

20.2:11

$c_1(x), c_2(x) \in \mathcal{I}$ ← both are inequality constraints

$$L(x,\lambda) = -2x_1 + x_2 - \lambda_1(1-x_1)^3 + \lambda_2 x_2 - \lambda_2 x_2 - \frac{1}{4}\lambda_2 x_1^2 + \lambda_2$$

$$\frac{\partial L}{\partial x_1} = -2 + 3\lambda_1(1-x_1)^2 - \frac{1}{2}\lambda_2 x_1, \quad \frac{\partial L}{\partial x_2} = 1 + \lambda_1 - \lambda_2$$

$$\nabla_x L(x^*, \lambda^*) = \begin{pmatrix} -2+3\lambda_1^*(1-x_1^*)^2 - \frac{1}{2}\lambda_2^* x_1^* \\ 1+\lambda_1^* - \lambda_2^* \end{pmatrix} = 0 \Rightarrow \lambda^* = \begin{pmatrix} \frac{2}{3} \\ \frac{5}{3} \end{pmatrix}$$

$$\hookrightarrow \begin{cases} -2+3\lambda_1 = 0 \\ 1+\lambda_1-\lambda_2 = 0 \end{cases} \quad \begin{cases} \lambda_1 = \frac{2}{3} \\ \lambda_2 = \frac{5}{3} \end{cases} \checkmark$$

All KKT conditions are satisfied

d. <u>Are the 2ND Order Necessary / sufficient conditions satisfied?</u>

$$H(L) = \nabla^2_{xx} L = \begin{pmatrix} -6\lambda_1^*(1-x_1^*) - \frac{1}{2}\lambda_2^* & 0 \\ 0 & 0 \end{pmatrix}$$

$$\hookrightarrow \frac{\partial^2 L}{\partial x_1^2} = -6x_1(1-x_2) - \frac{1}{2}\lambda_2, \quad \frac{\partial^2 L}{\partial x_1 \partial x_2} = 0, \quad \frac{\partial^2 L}{\partial x_2 \partial x_1} = 0, \quad \frac{\partial^2 L}{\partial x_2^2} = 0$$

$$= \begin{pmatrix} -\frac{29}{6} & 0 \\ 0 & 0 \end{pmatrix}$$ is semidefinite negative

$\nsucceq 0$!

But! From previous def: $\lambda_i^* = 0 \quad \forall i \in \mathcal{I} \setminus \mathcal{A}(x^*)$

Hint: λ_i^* is zero for inactive constraints. c_1, c_2 are both active, in fact $\lambda^* \neq 0$!

The only way to respect critical cone definition:

$$w \in \mathcal{C}(x^*, \lambda^*) \text{ s.t. } \underbrace{\lambda_i^*}_{\neq 0} \nabla c_i(x^*)^T w = 0 \quad \forall i \in \mathcal{E} \cup \mathcal{I}$$

is $w = (0,0)^T$ s.t. $w \in \mathcal{C}(x^*, \lambda^*)$

2^{ND} Order necessary conditions are satisfied

↳ KKT holds, $w^T H(L) w = 0$

↳ But sufficient are not

x^* could be a solution (week local min.) but it's not a strong local minimum

↳ Hint: we don't know, we have not verified it!

EX 4 Solve pb. 12.21 in the book. Illustrate the gradients of the active constraints and the objective function at the opt. point

---

4) Computing Solution

- Linear Programming (LP)
  - Simplex Method
  - Interior-Point Methods
- Non Linear Programming (NLP)
  - Penalty Methods
  - Sequential Quadratic Programming Methods
  - Interior-Point Methods

Hint: solve a QP problem at each iterate