Basic Concepts in Optimization

Prof. Cesar de Prada
Dpt. Systems Engineering and Automatic Control
UVA
prada@autom.uva.es

Optimization problems

$$\min_{x} J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \leq 0$$

NPL problem

In order to find the solution of these problems, it is important:

1. Analyse the properties of its mathematical expressions
2. Analyse the mathematical structure of the problem, classify it according to this structure and find appropriate methods for each of them.

Outline

- General concepts
  - Formulation
  - Local and global optimum
  - Feasibility
- Mathematical properties
  - Continuity
  - Convexity
- Different types of optimization problems

Terminology

$$\min_{x} J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \leq 0$$

$$x \in R^n$$

$x = (x_1, x_2, ..., x_n)'$ decision vector of real variables

$J(x)$ cost function

$h_i(x) = 0 \quad i = 1,2,...,l$ equality constraints

$g_j(x) \leq 0 \quad j = 1,2,...,m$ inequality constraints

If $h_i(x)$ and $g_j(x)$ do not exist, the problem is called unconstraint optimization

Equivalencies

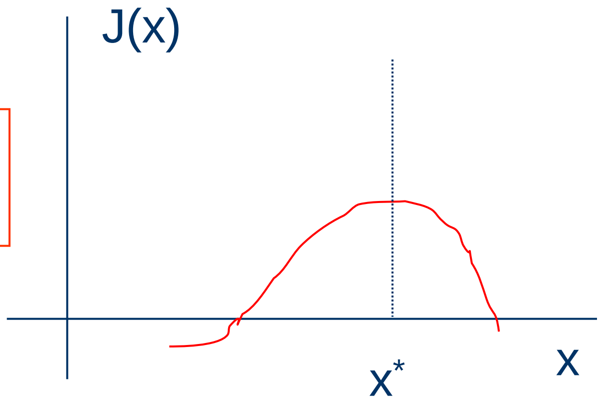
$$\min_{x} J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \le 0$$

Minimize / Maximize

min J(x) = max –J(x)

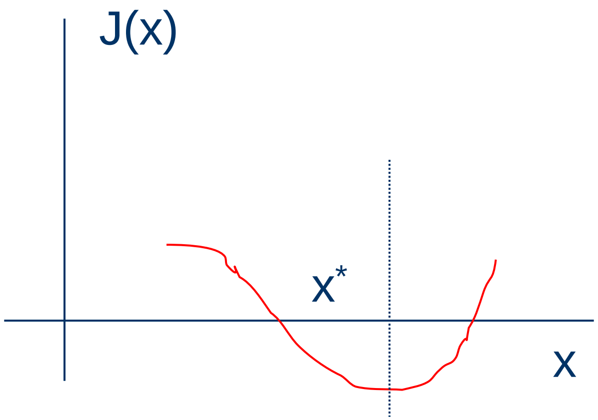


$g_j(x) \le a$ can be written as $g_j(x) - a \le 0$

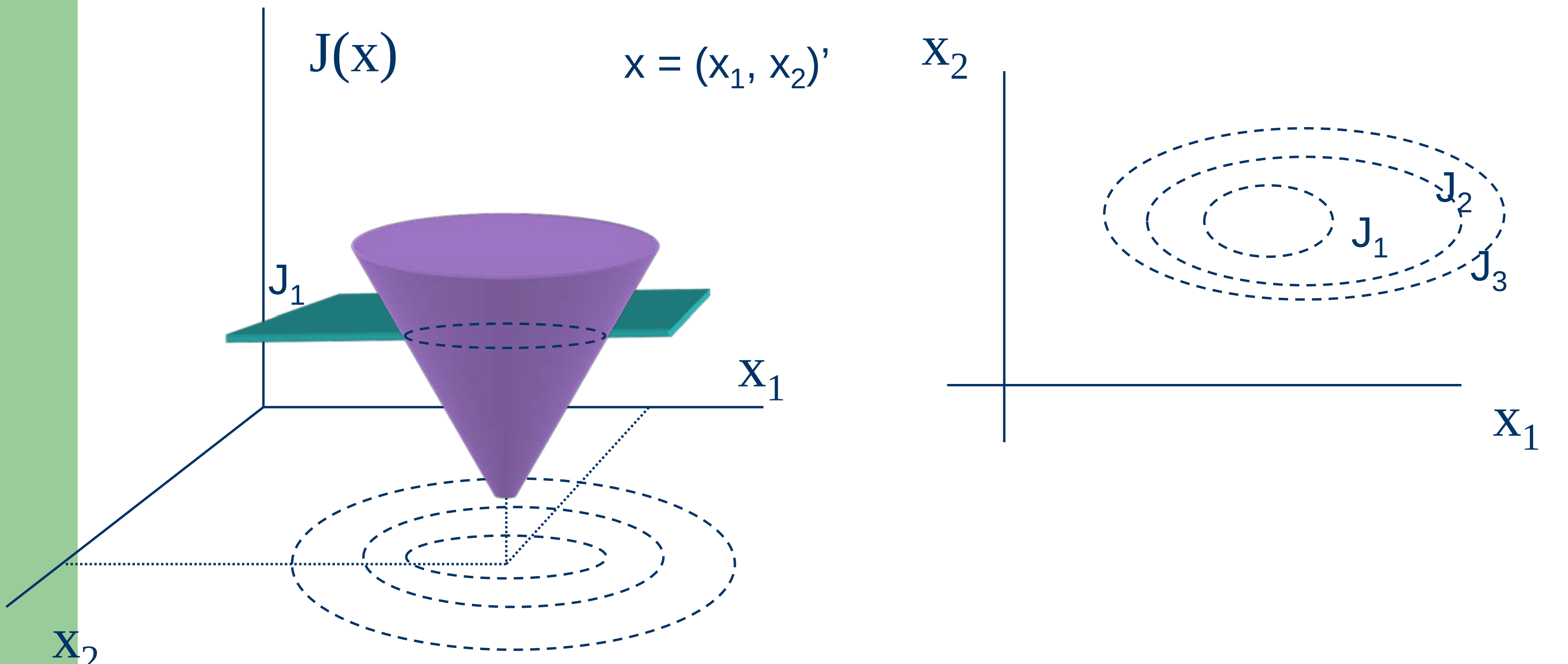
$g_j(x) \le 0$ is equivalent to $-g_j(x) \ge 0$

$g_j(x) \le 0$ is equivalent to $g_j(x) + \varepsilon = 0,\ \varepsilon \ge 0$

$h_i(x) = 0$ is equivalent to $h_i(x) - \varepsilon \le 0,\ \varepsilon \ge 0$



Contours



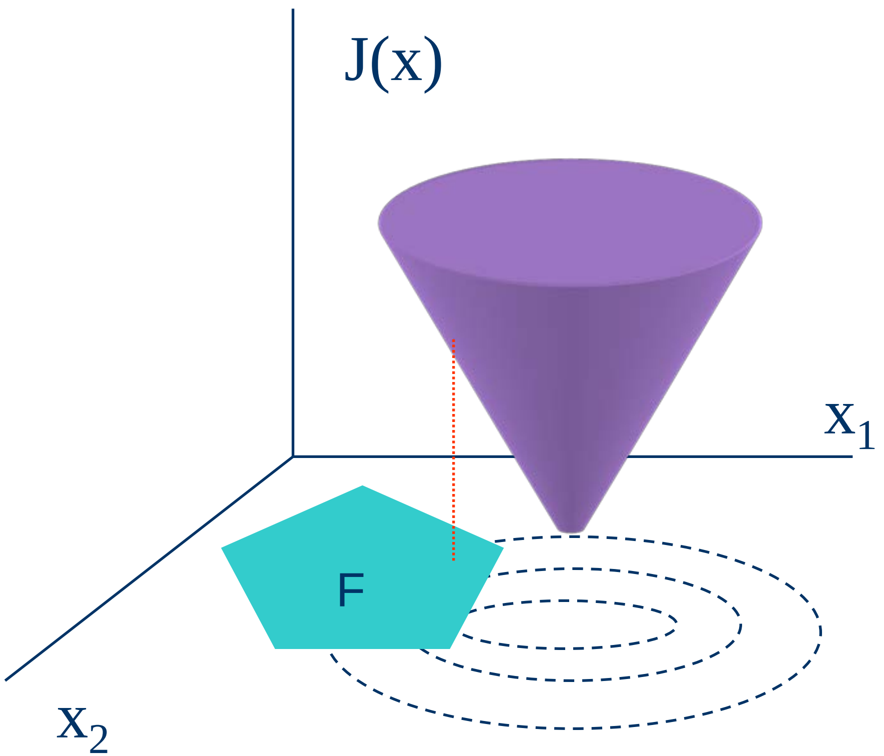
Feasibility

$$\min_x J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \le 0$$

The constraints define the searching space or feasible set F

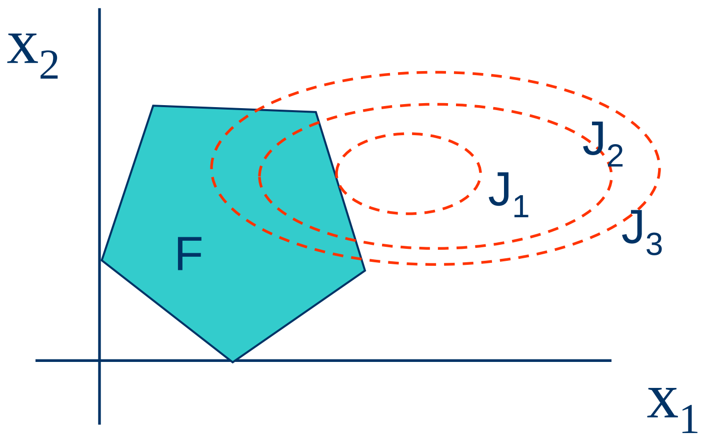


Feasibility

$$\min_{x} J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \leq 0$$



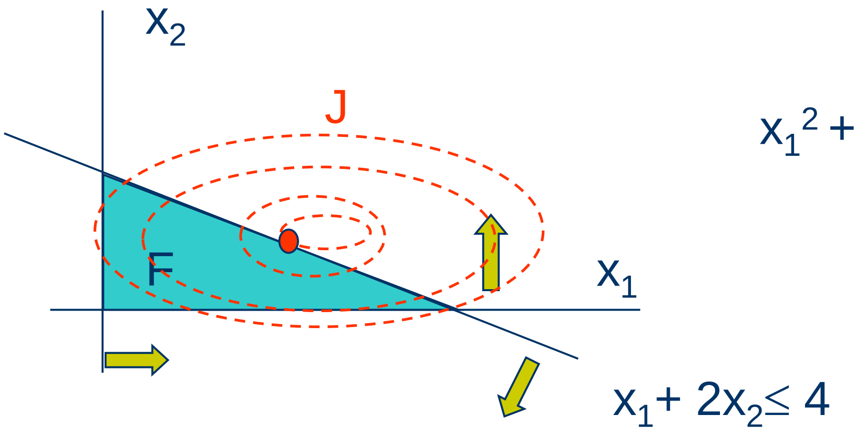
The constraints define the searching space or feasible set F

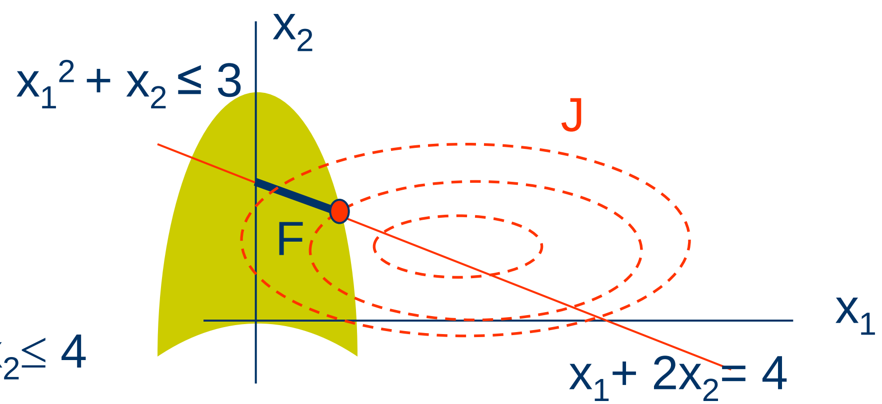
If the set F is empty, that is, there is no x satisfying all constraints, the problem is no feasible and it has no solution

Examples

$$\min\ (x_1 - 2)^2 + 3(x_2 - 1)^2 + 1$$
$$x_1 + 2x_2 \le 4$$
$$x_1 \ge 0$$
$$x_2 \ge 0$$

$$\min\ (x_1 - 2)^2 + 3(x_2 + 1)^2 + 1$$
$$x_1 + 2x_2 = 4$$
$$x_1^2 + x_2 \le 3$$
$$x_1 \ge 0$$
$$x_2 \ge 0$$





Active constraints

$$\min_{x} J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \le 0$$

$$\min\,(x_1 + 3x_2^2)$$

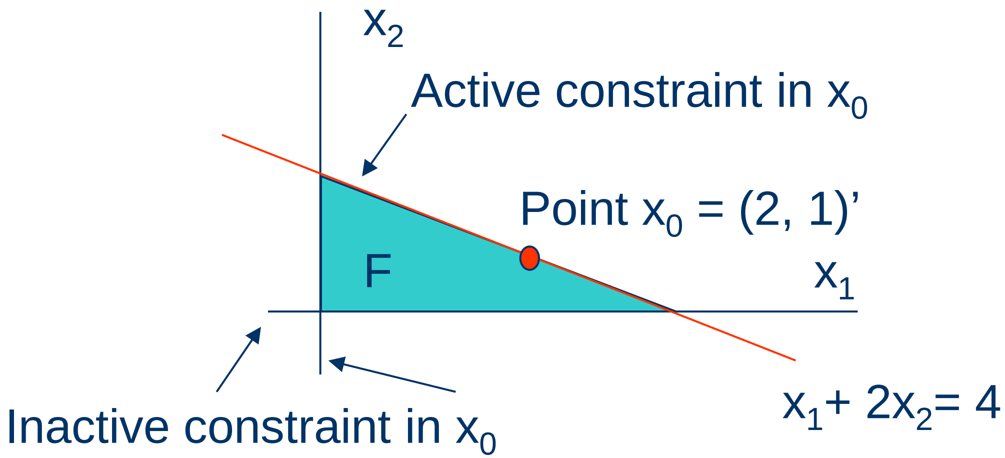
$$x_1 + 2x_2 \le 4$$

$$x_1 \ge 0$$

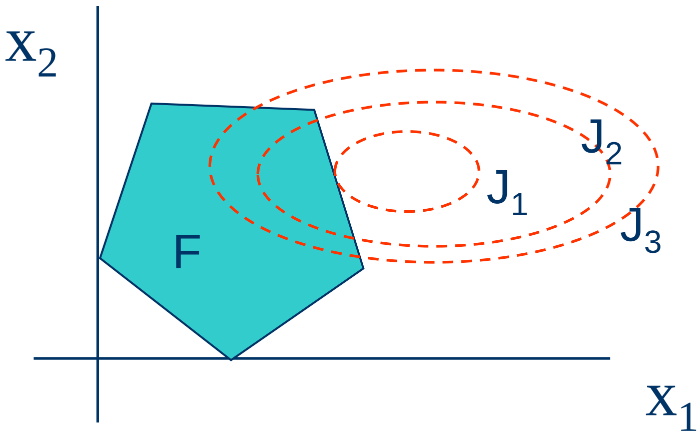
$$x_2 \ge 0$$

A constraint $g_j(x) \le 0$ is active in a point x_0 if: $g_j(x_0) = 0$

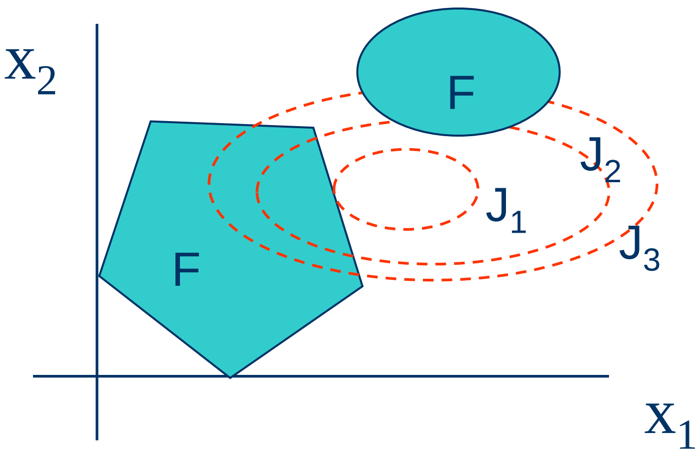
(Quite often the concept refers to the optimal solution)



Conex regions



Conex Feasible region F

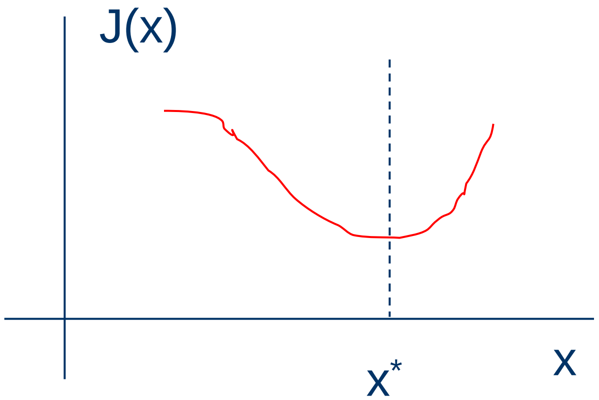


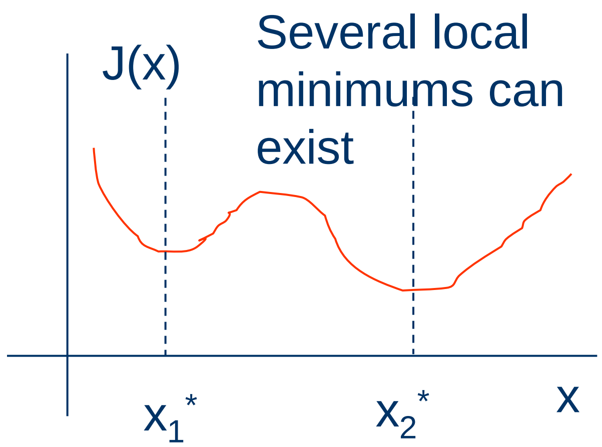
Non conex feasible region F

Local optimum (local)

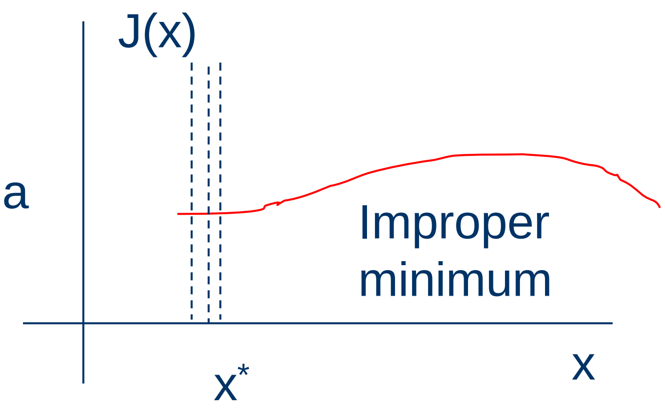
A point $x^* \in F$ is call a local minimum of the optimization problem if there exist a set around x^* such that for any other point $x \in F$ from the set:

$$J(x^*) \leq J(x)$$





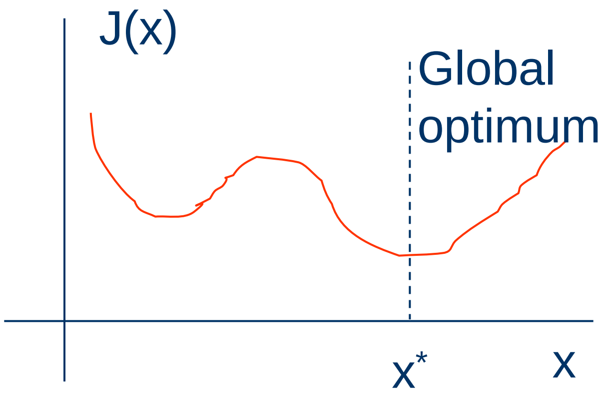
If the inequality is strict, the minimum is a proper one

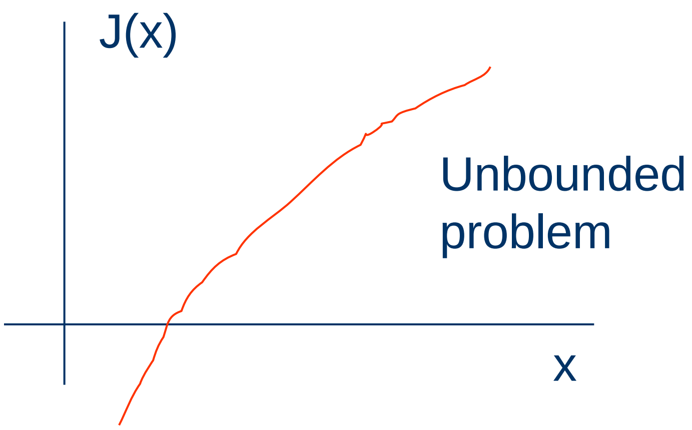


Global optimum

A point x^* is called a global optimum of the optimization problem if for any point belonging to the feasible set F:

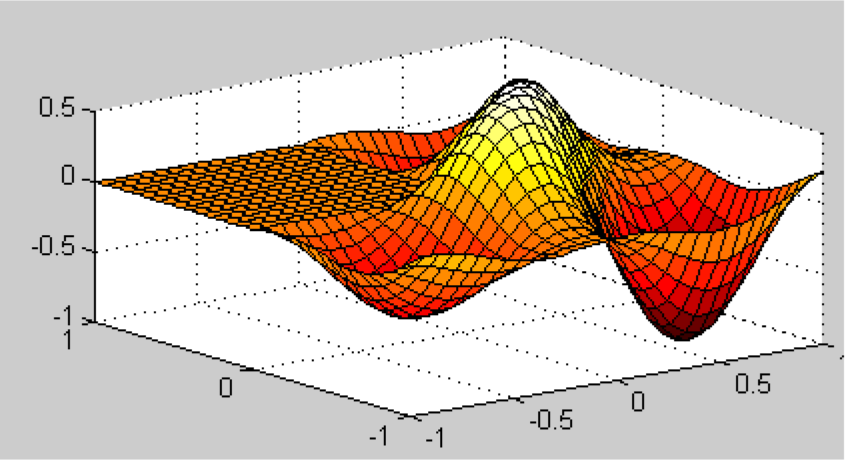
$$J(x^*) \le J(x)$$



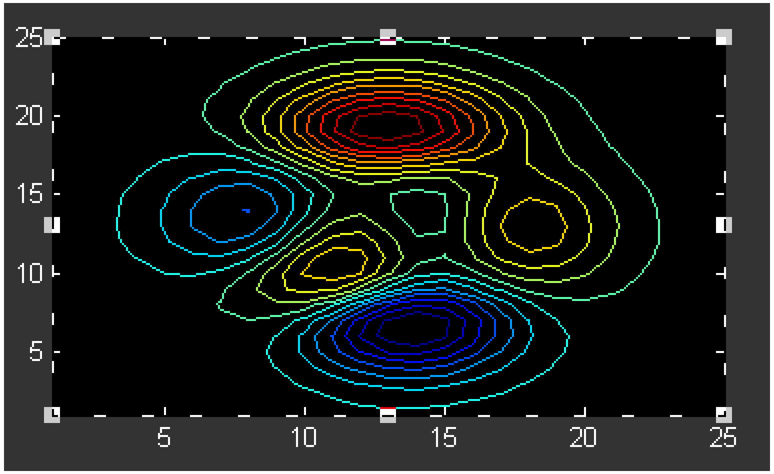


If there is no $x^* \in F$ such that $J(x^*) \le J(x)$ then the problem is unbounded and there is no minimum

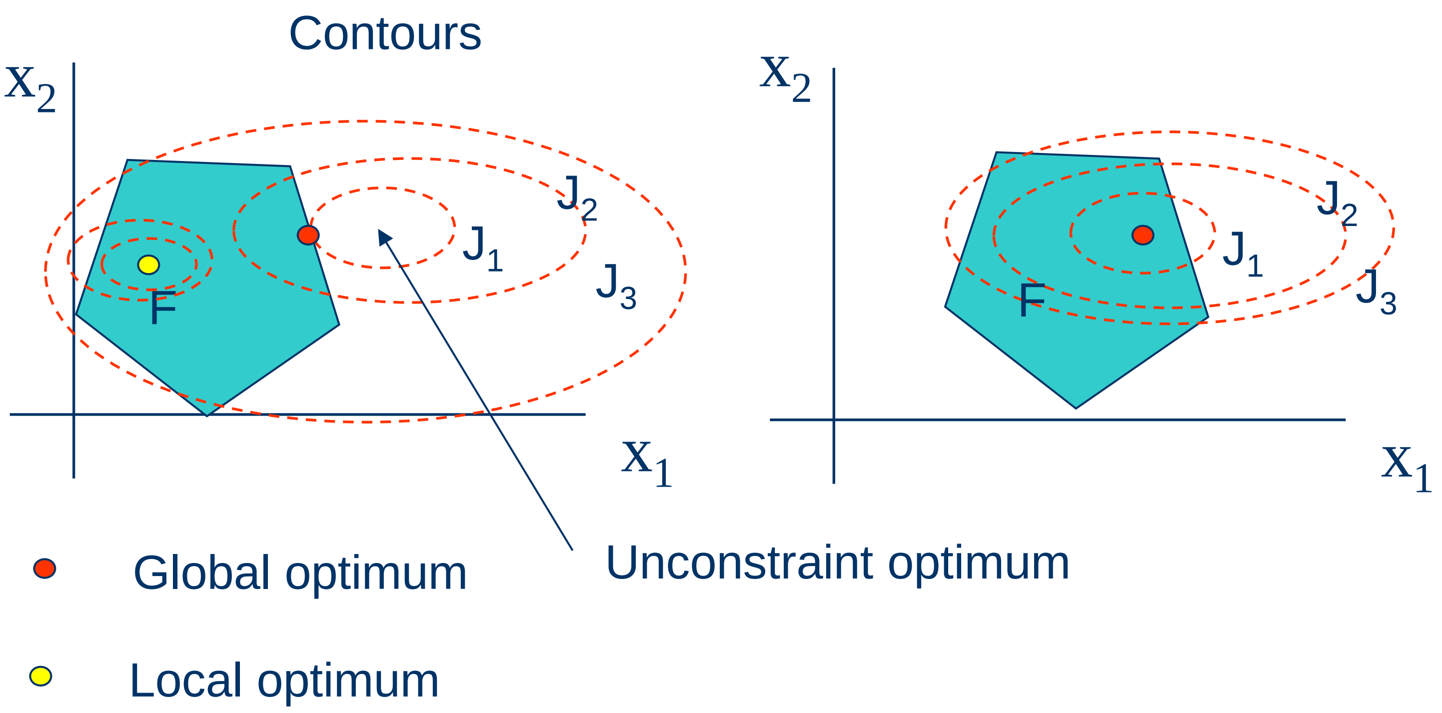
Example



Several local minimums and maximums

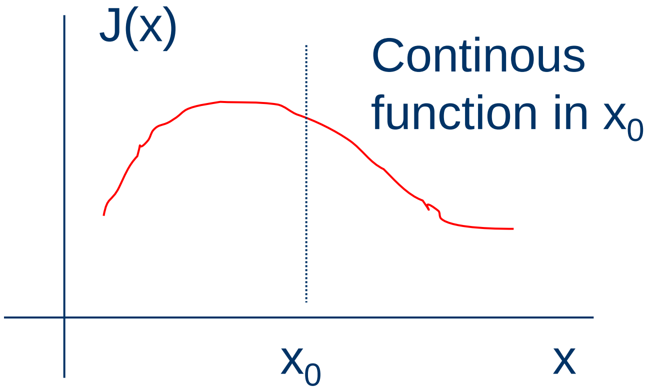


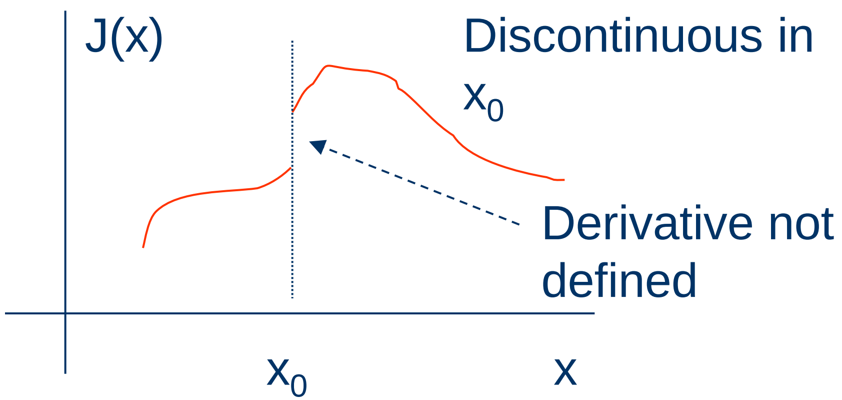
Examples



Contours
x2
J2
J1
J3
F
x1
x2
J2
J1
J3
F
x1
Global optimum
Unconstraint optimum
Local optimum

Continuity





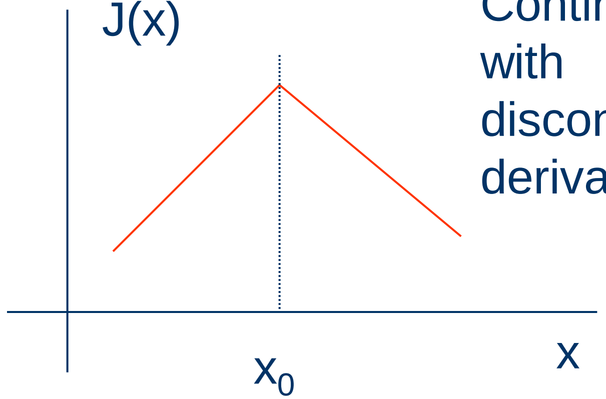
$$\lim_{x \to x_0} J(x) \quad \text{exist}$$

$$J(x_0) \quad \text{exist}$$

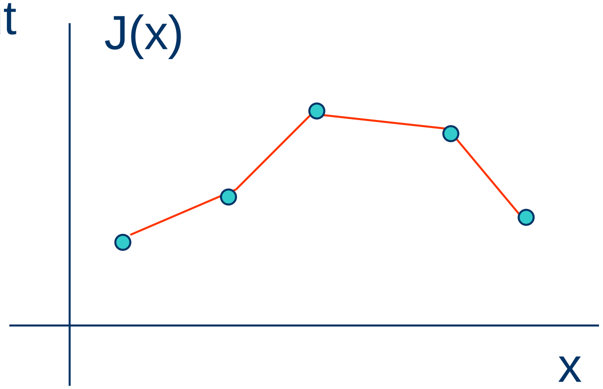
$$\lim_{x \to x_0} J(x) = J(x_0)$$

Many algorithms require continuous functions and continuous derivatives

Continuity



Continuous, but with discontinuous derivative in x_0

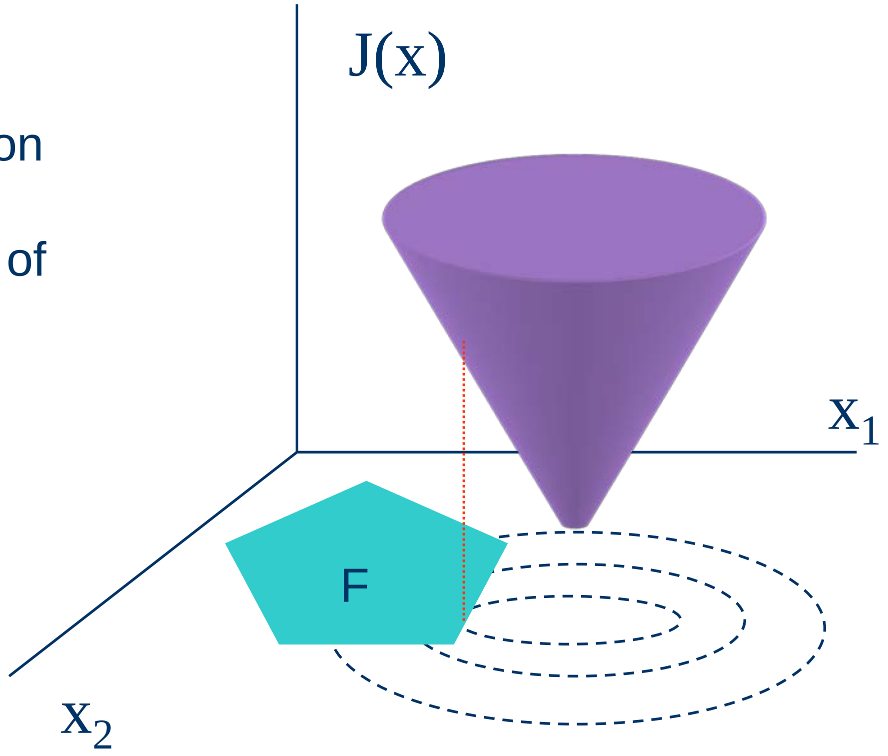


Those optimization methods based on the use of derivatives can suffer from oscillations and lack of convergence if there are discontinuities in the functions

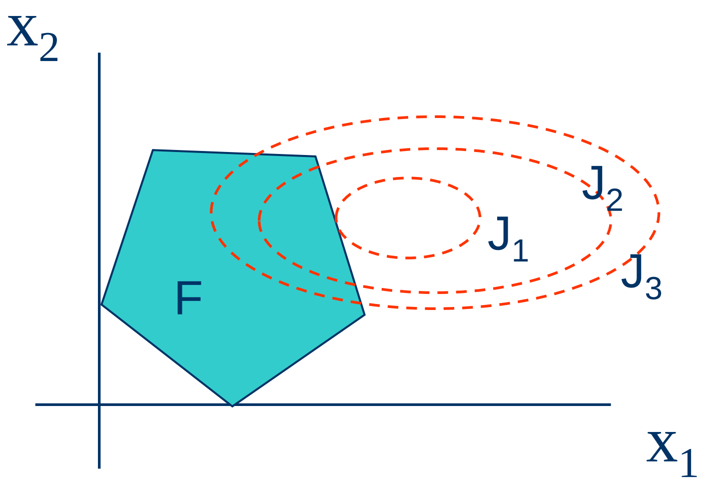
Discontinuous derivatives appear when linear interpolation is used to compute values of a function defined only at a discrete number of points x

Theorem

A continuous function J(x) has a global minimum at a point of any closed and bounded set F

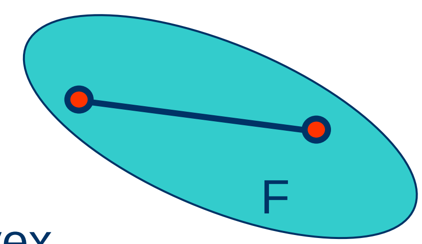


Convexity



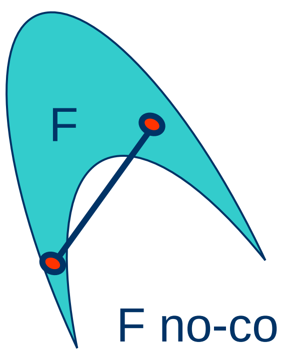
The shape of the searching area is important for the optimization methods

$$\min_{x} J(x)$$
$$h_i(x) = 0$$
$$g_j(x) \leq 0$$



F convex

A set F is a convex one if and only if, the segment joining any two points of the set is fully included in the set F



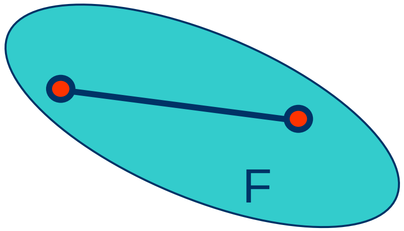
F no-convex

Convex set

F is convex if, and only if:

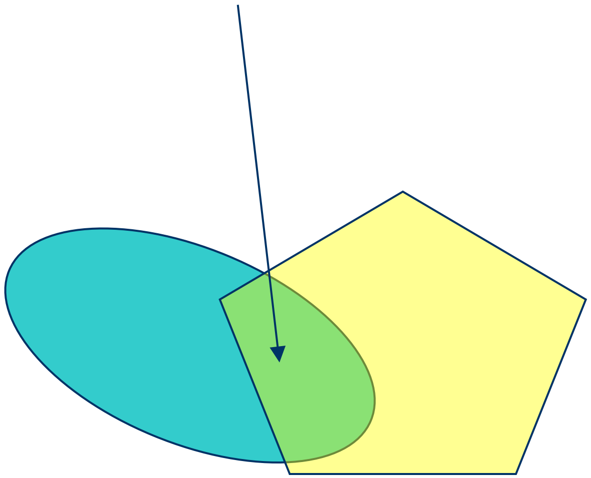
$$\forall x_1, x_2 \in F, \quad \forall \gamma \in [0,1]$$

$$x = \gamma x_1 + (1-\gamma) x_2 \in F$$



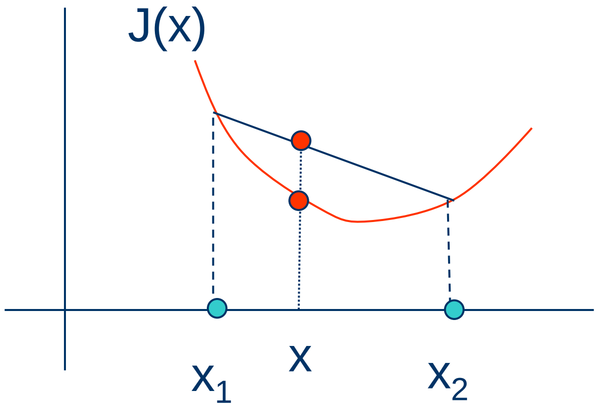
Closed and convex region

The intersection of two convex sets is convex



Convex functions

Function J(x) is convex in a convex set F if it is always bellow a linear interpolation between any two points



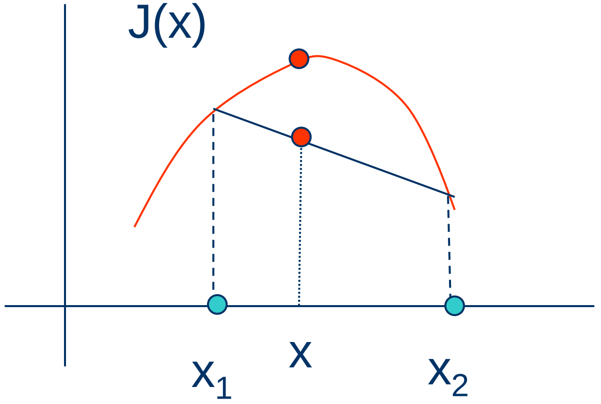
$$\forall x_1, x_2 \in F, \quad \forall \gamma \in [0,1]$$

$$J(\gamma x_1 + (1-\gamma) x_2) \leq \gamma J(x_1) + (1-\gamma) J(x_2)$$

If the inequality stands with < the function is strictly convex

Concave functions

Function J(x) is concave in a convex set F if it is always above a linear interpolation between any two points

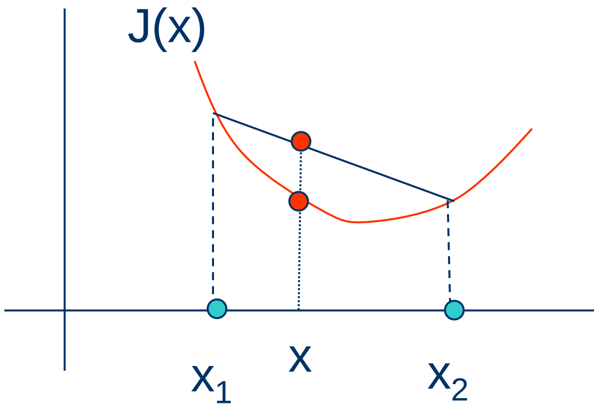


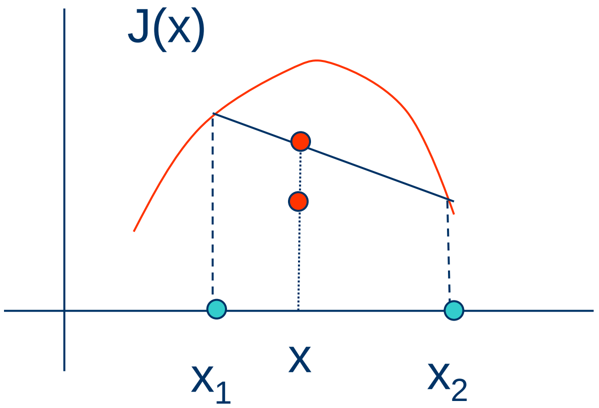
$$\forall x_1, x_2 \in F, \quad \forall \gamma \in [0,1]$$

$$J(\gamma x_1 + (1-\gamma)x_2) \geq \gamma J(x_1) + (1-\gamma)J(x_2)$$

If the inequality stands with > the function is strictly concave

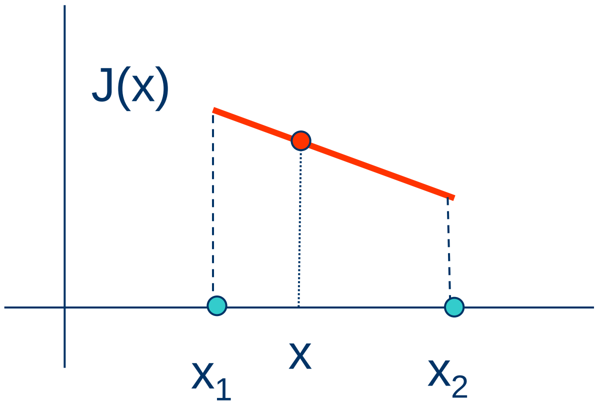
Convexity



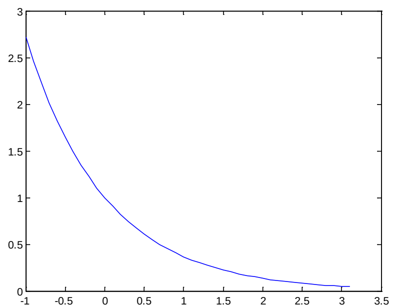


If J(x) is convex then -J(x) is concave

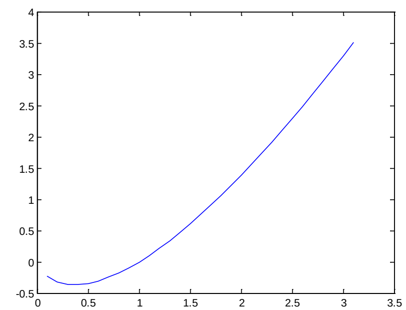
A linear function is convex and concave



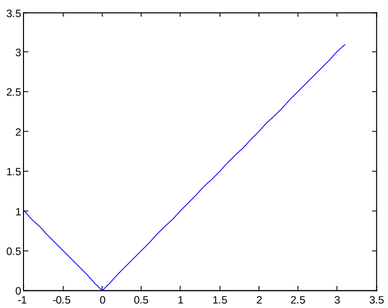
Examples of convex functions



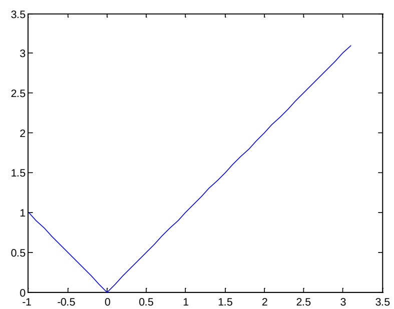
exp(-x)



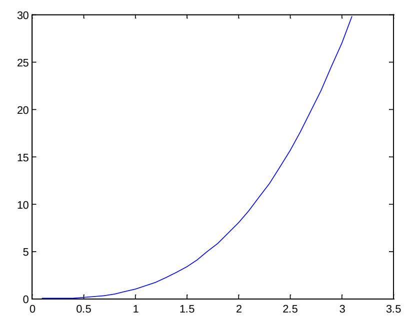
x log(x)   x > 0



$\sigma_{max}(x)$



|x|



x^a $a \geq 1, x > 0$

All norms are convex. The geometric mean is concave

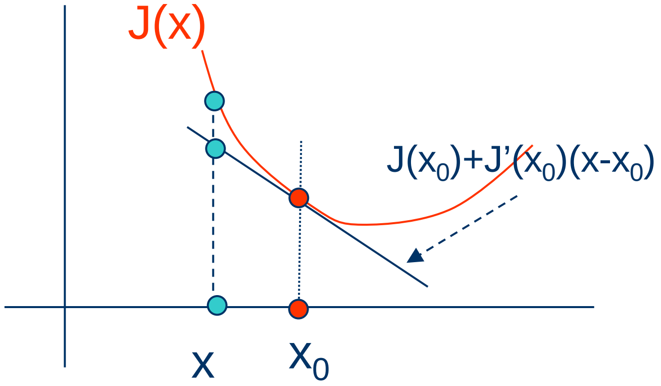
(Local) Convexity of one variable functions

$$J(x) = J(x_0) + \frac{dJ(x_0)}{dx}(x - x_0) + \frac{1}{2}\frac{d^2J(x_0)}{dx^2}(x - x_0)^2 + \ldots$$

$$J(x) - (J(x_0) + \frac{dJ(x_0)}{dx}(x - x_0)) = \frac{1}{2}\frac{d^2J(x_0)}{dx^2}(x - x_0)^2 + \ldots$$

$$H = \frac{d^2J(x_0)}{dx^2}$$

If H is continuous and positive semidefinite, then J(x) is convex in an interval around x_0



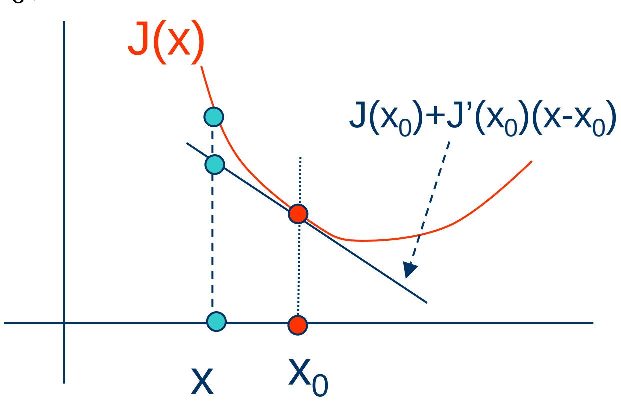
(Local) Convexity of functions

$$J(x) = J(x_0) + \left.\frac{\partial J}{\partial x}\right|_{x_0} (x - x_0) + \frac{1}{2}(x - x_0)' \left.\frac{\partial^2 J(x)}{\partial x^2}\right|_{x_0} (x - x_0) + ..$$

$$J(x) - (J(x_0) + \left.\frac{\partial J}{\partial x}\right|_{x_0} (x - x_0)) = \frac{1}{2}(x - x_0)' \left.\frac{\partial^2 J(x)}{\partial x^2}\right|_{x_0} (x - x_0) + ..$$

$$\frac{1}{2}(x - x_0)' \left.\frac{\partial^2 J(x)}{\partial x^2}\right|_{x_0} (x - x_0) = \frac{1}{2}(x - x_0)' H (x - x_0)$$

$\partial J/\partial x$ Jacobian
H Hessian



The quadratic form z'Hz defines the convexity of J(x) around x_0

Quadratic forms / PD matrices

A quadratic form z'Hz is positive definite (PD) if z'Hz > 0 $\forall z$

The matrix H must have all its eigenvalues > 0

By extension, H is named also as PD

A quadratic form z'Hz is positive semidefinite (PSD) if z'Hz ≥ 0 $\forall z$

The matrix H must have all its eigenvalues ≥ 0

A quadratic form z'Hz is negative definite (ND) if z'Hz < 0 $\forall z$

The matrix H must have all its eigenvalues < 0

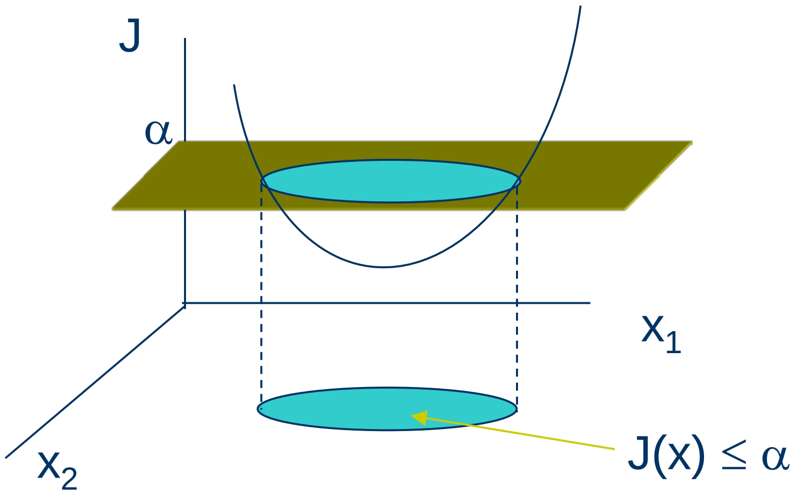
A quadratic form z'Hz is indefinite if z'Hz can have positive and negative values

The matrix H must have positive and negative eigenvalues

Region $J(x) \le \alpha$

If the function J(x) is convex in a convex set F, then the set:

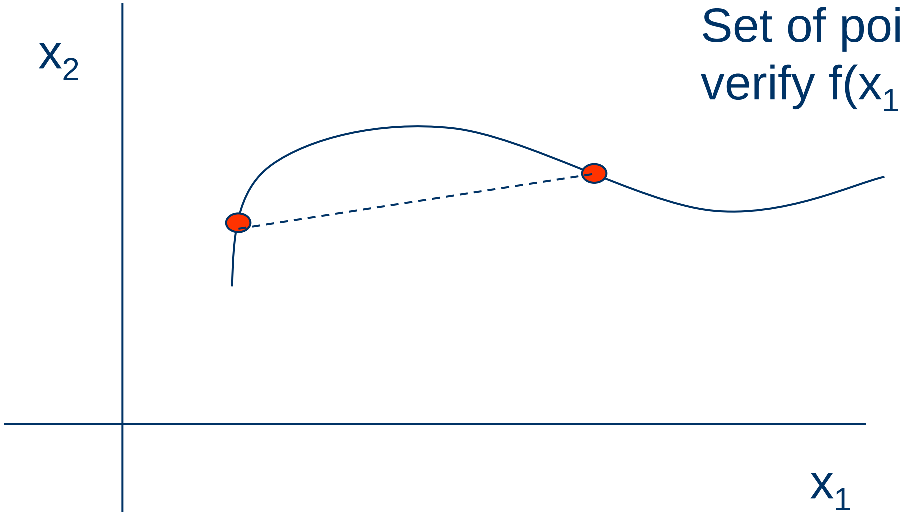
$$\left\{ x \mid x \in F, J(x) \le \alpha \right\}$$ is convex



Set f(x)=0

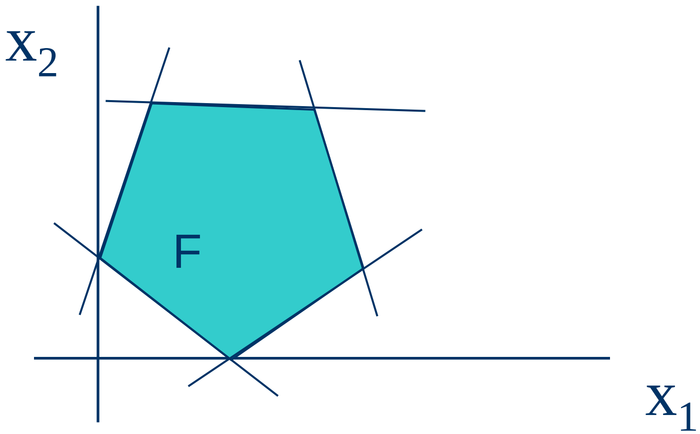
In general, a set of points x defined by f(x) = 0 is non convex

Set of points that verify $f(x_1,x_2) = 0$

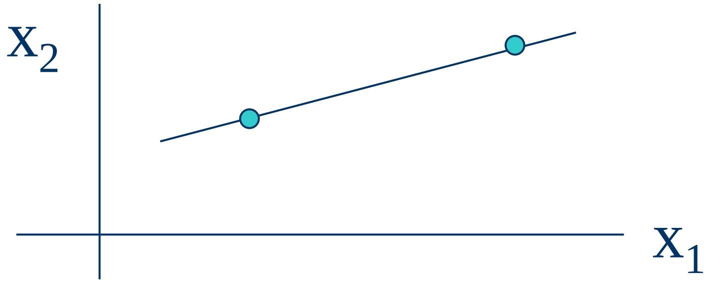


Convexity of linear functions

Regions defined by linear inequalities are convex. They are called polytopes. A bounded polytope is called a polyhedron



Linear functions are convex (and concave)



Quadratic functions

$$J(x) = a + b'x + \frac{1}{2} x'Hx$$

$$\frac{\partial J(x)}{\partial x} = b' + x'H$$

$$\frac{\partial^2 J(x)}{\partial x^2} = H$$

Matrix H defines the convexity of the function

The convexity is global

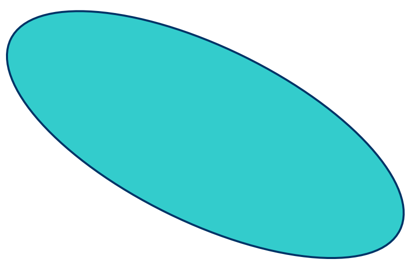
$$J(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Quadratic function (form) in R^2

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq 1$$

Describes a set in R^2

Convexity of quadratic regions



The set $x'Hx \leq 1$ is convex if the matrix H is real simetric positive semidefinite

H is positive semidefinite if $Q(x) = x'Hx \geq 0 \quad \forall x \neq 0$, eigenvalues ≥ 0

H es positive definite if $Q(x) = x'Hx > 0 \quad \forall x \neq 0$, eigenvalues > 0

H es negative semidefinite if $Q(x) = x'Hx \leq 0 \quad \forall x \neq 0$, eigenvalues ≤ 0

H es negative definite if $Q(x) = x'Hx < 0 \quad \forall x \neq 0$, eigenvalues < 0

The quadratic function Q(x) is PD if H is PD, etc.

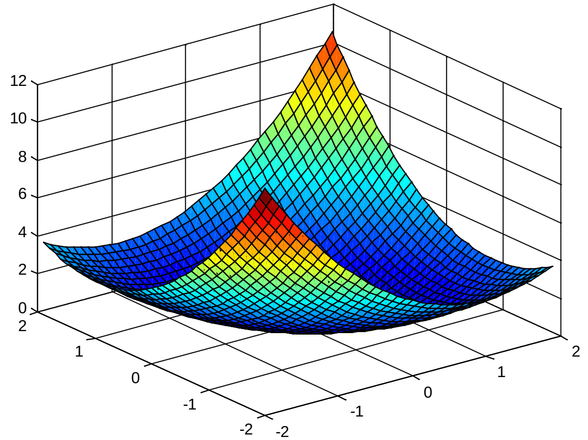
Example

Function

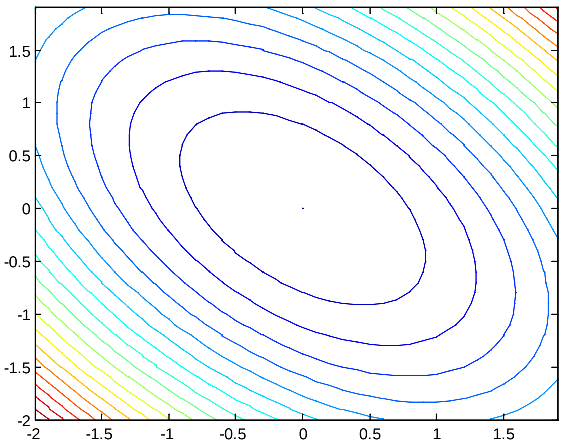
$$J(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Set

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \alpha$$



Eigenvalues 1.5, 0.5 PD



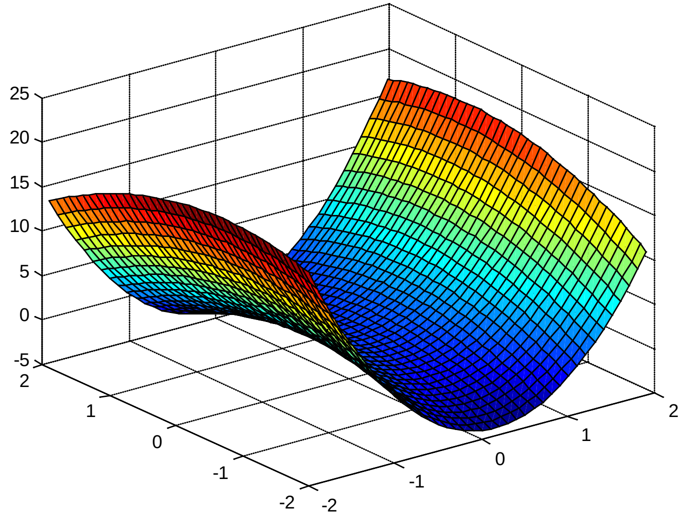
Contours $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha$

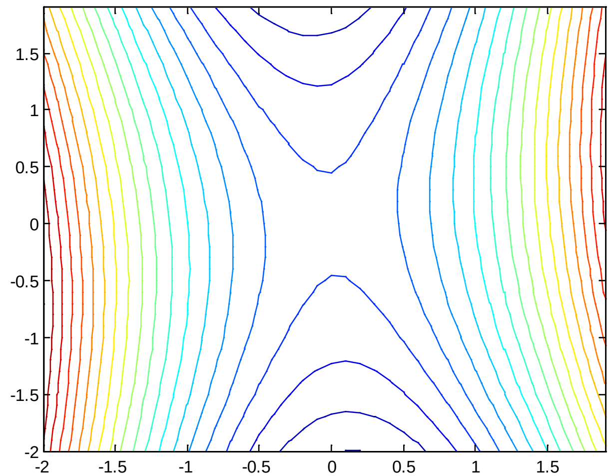
Eigenvalues 5.02, -2.02
Indefinite

Example

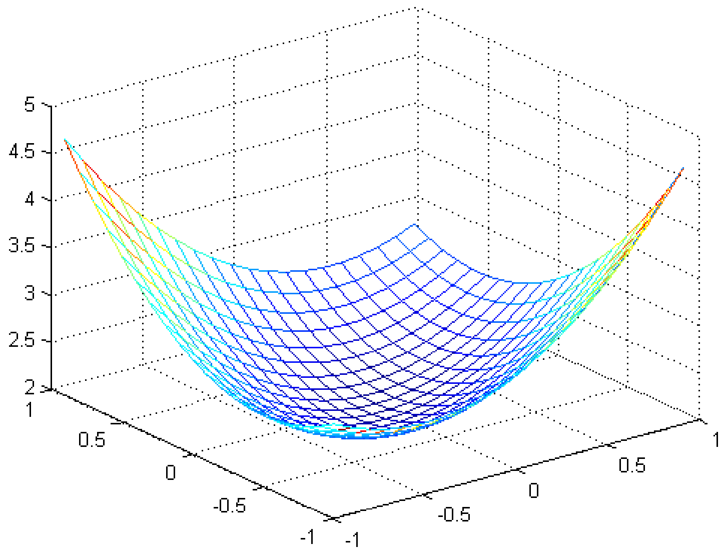
$$J(x_1,x_2)=\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}5 & 0.5\\ 0.25 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

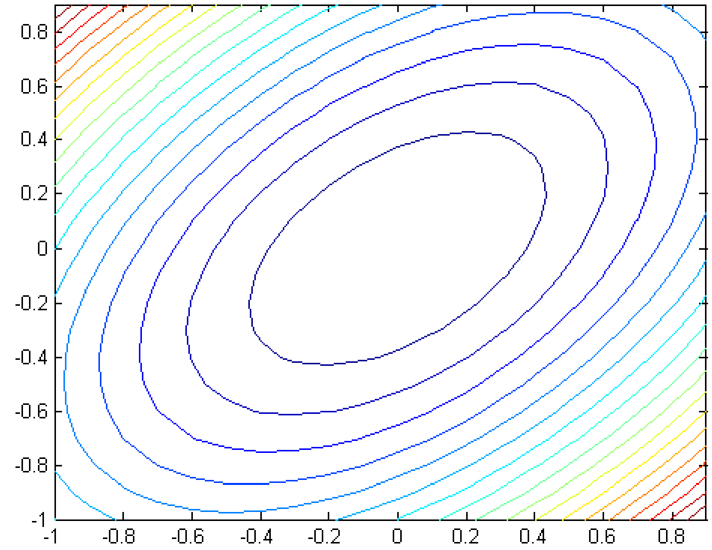
$$\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}5 & 0.5\\ 0.25 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}\le 1$$





PD
Quadratic function



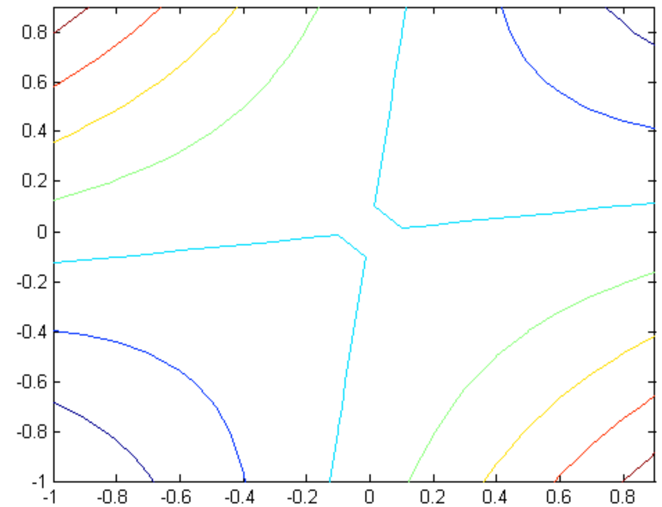


$$J(x_1, x_2) = x_1^2 + x_2^2 - x_1 x_2 + 2$$

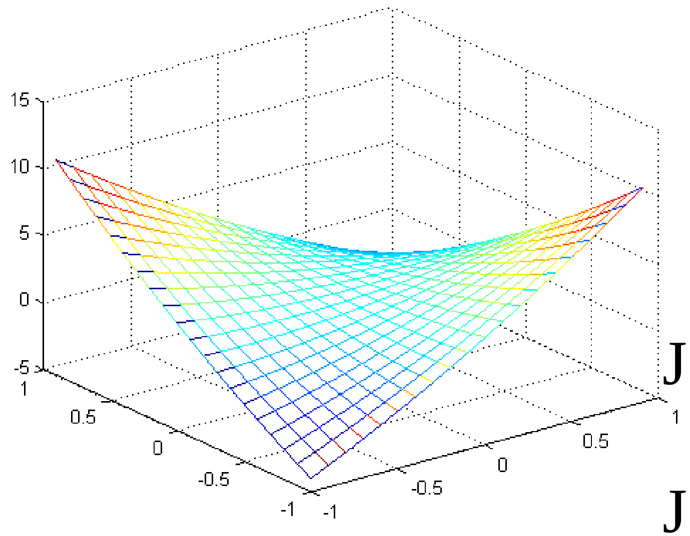
$$J(x_1, x_2) = \frac{1}{2}(x_1, x_2)' \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 2$$

$$\frac{\partial^2 J}{\partial x^2} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \mathrm{eig}\left[\frac{\partial^2 J}{\partial x^2}\right] = 1, 3$$

Indefinite quadratic function



Saddle point



$$J(x_1, x_2) = x_1^2 + x_2^2 - 8x_1x_2 + 2$$

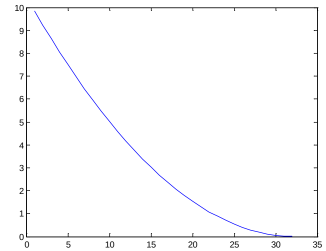
$$J(x_1, x_2) = \frac{1}{2}(x_1, x_2)'\begin{bmatrix} 2 & -8 \\ -8 & 2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 2$$

$$\frac{\partial^2 J}{\partial x^2} = \begin{bmatrix} 2 & -8 \\ -8 & 2 \end{bmatrix} \quad \mathrm{eig}\left[\frac{\partial^2 J}{\partial x^2}\right] = -6, 10$$

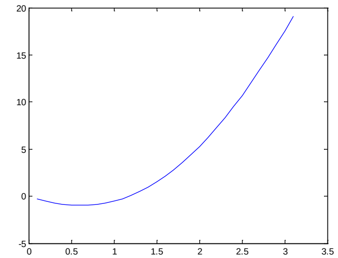
Convexity of general functions

- ✓If $J_1(x)$ and $J_2(x)$ are convex functions in the convex set F, then $J_1(x) + J_2(x)$ is also convex in F
- ✓If $J_1(x)$ and $J_2(x)$ are convex functions with an upper bound in the convex set F, then $J(x) = \max \{ J_1(x), J_2(x)\}$ is also convex in F
- ✓If $J_1(x)$ y $J_2(x)$ are concave functions with a lower bound in the convex set F, then $J(x) = \min \{ J_1(x), J_2(x)\}$ is also concave in F
- ✓If J(x) is convex in the convex set F, then J(Ax+b) is convex
- ✓If J(x) is a convex function in the convex set F, and if V(.) is a non decreasing convex function (defined in the range of J), then V[J(x)] is also convex F. This is also true if J(x) is concave and V is convex and non increasing

Convexity

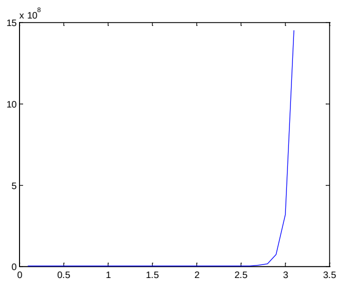


$(x - \pi)^2$

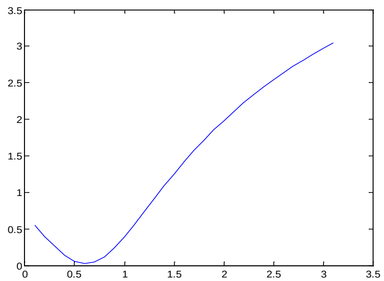


$2x^2 - 3\text{sen}(x)$

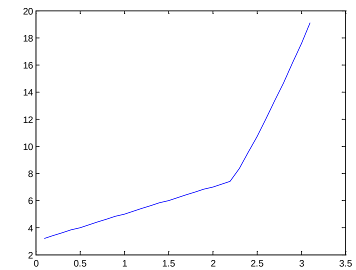
Analyse the convexity of … in the interval $(0,\pi]$



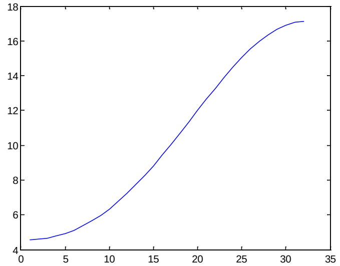
$\exp(2x^2 - 3\text{sen}(x) + 2)$



$\log(2x^2 - 3\text{sen}(x) + 2)$

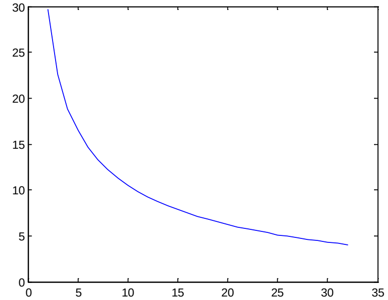


$\max\left\{\begin{matrix} 2x + 3, \\ 2x^2 - 3\text{sen}(x) \end{matrix}\right\}$

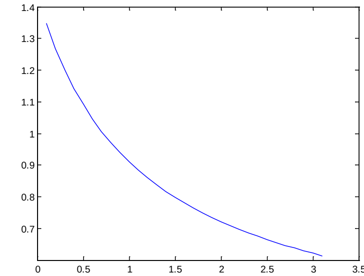


$(\cos x - \pi)^2$

$x \in (0,\pi]$

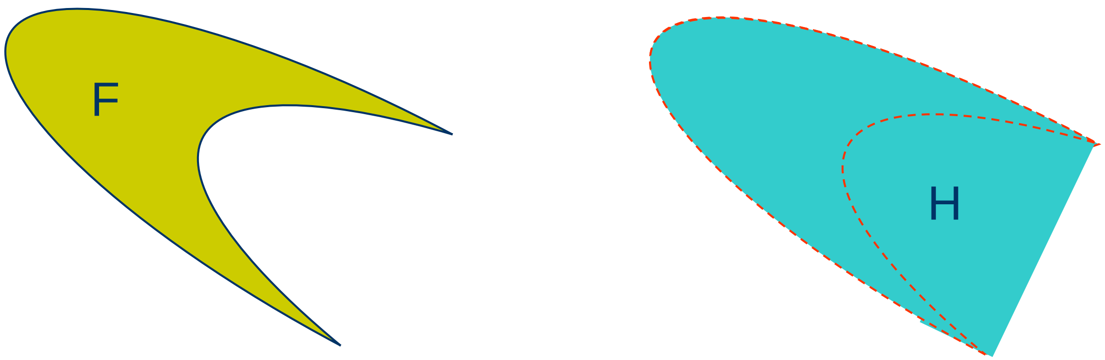


$(\log x - \pi)^2$



$\dfrac{1}{\log x + 2}$

Convex hull



The convex hull of F is the minimum convex set containing F

Summary

- The convexity of a function at a point x can be studied by means of its Hessian H
- A function with continuous hessian H, defined in a convex set F (with at least an interior point) is convex if, and only if, H is a positive semidefinite matrix in F.
- The set F defined by the expressions $g_j(x) \leq 0$ and $h_i(x)=0$ is convex si all g_j are convex and all h_i are linear

Optimization in a convex set

$$\min_x J(x)$$

$$h_i(x) = 0$$

$$g_j(x) \leq 0$$

If J is convex in the convex set F, then a local minimum is also a global one.

If all inequality constraints are convex, they will generate a convex feasible set F. If any equality constraints is non-linear, it will not be convex, hence the problem could have local minimums.

Different types of optimization problems

$$\min_x J(x)$$
$$x \in R^n$$

Unconstraint optimization

$$\min_x J(x)$$
$$h_i(x) = 0$$

Optimization with equality constraints

Lagrange multipliers

Different types of optimization problems

$$\min_x b' x$$
$$Ax \leq c$$
$$x \geq 0$$

Linear Programming (LP)
The cost function and the constraints are linear

$$\min_x x' Hx + b' x$$
$$Ax \leq c$$
$$x \geq 0$$

Quadratic Programming (QP)
The cost function is quadratic and the constraints are linear

Different types of optimization problems

$$\min_x J(x)$$
$$h_i(x) = 0$$
$$g_j(x) \le 0$$

Non Linear Programming (NLP)
The cost function or some constraints are non linear

$$\min_x J(x, y)$$
$$h_i(x, y) = 0$$
$$g_j(x, y) \le 0$$
$$x \in R^n, y \in Z$$

Mix Integer Programming (MINLP)
Some of the variables are integers and other are real

Different types of optimization problems

$$\min_{x} J(x,z)$$

$$\frac{dz}{dt} = f(z,x)$$

$$g_j(x) \leq 0$$

$$r_i(z) \leq 0$$

Dynamic Optimization

Some of the constraints are given as differential equations

$$\min_{x} \{J_1(x), J_2(x), ...J_s(x)\}$$

$$x \in \Omega$$

Multiobjective Optimization
There are several cost functions to be optimized simultaneously