INTRODUCTION

Notations :

$x(t) \in \mathbb{R}^n$ state variable

$u(t) \in \mathbb{R}^p$ control variable

$$f: \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$$

$\bar{C}^k$ function with a derivative up to the k-th order continuous almost everywhere

example:

$\bar{C}^0 \rightsquigarrow$ continuous almost everywhere

$\bar{C}^1 \rightsquigarrow$ continuous (C^0) and C^1 almost everywhere

$\vdots$

$\bar{C}^6 \rightsquigarrow$ continuous $C^0, C^1, \dots, C^5$ and C^6 almost everywhere

Optimal control sets out to provide analytical designs of a special appealing type.

The final system is suppose to be the best possible system of a particular type

There is a COST INDEX because the system has to behave in the best way

best depends on the specific criteria chosen

- **Linear Optimal Control (LOC)**

special sort of optimal control

- → linear plant
- → controller constrained to be linear

Linear controllers are achieved thanks to <u>quadratic cost indices</u>

Pros:

1) LOC may be applied to nonlinear systems
2) LOC have mainly computational solutions
3) If a nonlinear system has not strong nonlinearities it is possible to model it approximating as a linear system

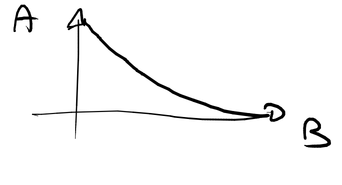
Birth of Optimal Control (1696)

↓

Bernoulli - Brachistochrone problem

Control the path of the behaviour of a dynamical system:

Particle of mass M moves along a wire from A to B under gravity. Find the shape of the wire in order to reach B in minimum time.



- **Feedback**

The actual operation of the control system is compared to the desired operation and the input to the plant is adjusted on the basis of this comparison

Feedback control systems are able to operate satisfactorily despite adverse conditions, such as disturbances and variations in plant properties

- **Optimal control problem**

$$\left.\begin{aligned} \dot{x} &= f(x,u,t) \\ x(t_0) &= x_0 \end{aligned}\right\} \text{ control system}$$

$$J(u) = \int_{t_0}^{t_f} L(x(t), u(t), t)\,dt + K(x(t_f), t_f)$$

↳ the cost index depends on the goal chosen

Remark: the optimal control $u(t)$ is a function of time.
J is a functional (a function of a function) which returns a number (cost)

The problem is to choose the best path among all paths feasible for the system, with respect to the given cost function.

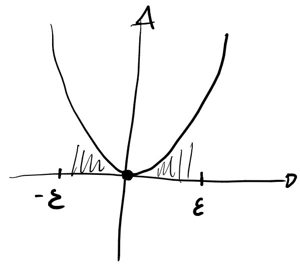
This is an infinite dimensional problem because the space of paths is an infinite dimensional function space

The problem is a dynamic optimization because it involves a dynamical system and time

- **local / strict / global minimum (maximum)**

Find a minimum of $f : \mathbb{R}^n \to \mathbb{R}$

example:



$y = x^2$

$y' = 2x \leadsto y' = 0 \to 2x = 0 \to x = 0$

$y'' = 2 > 0$

$f : \mathbb{R}^n : \to \mathbb{R}$ $\quad\quad x = (x_1 \dots x_n)^T$

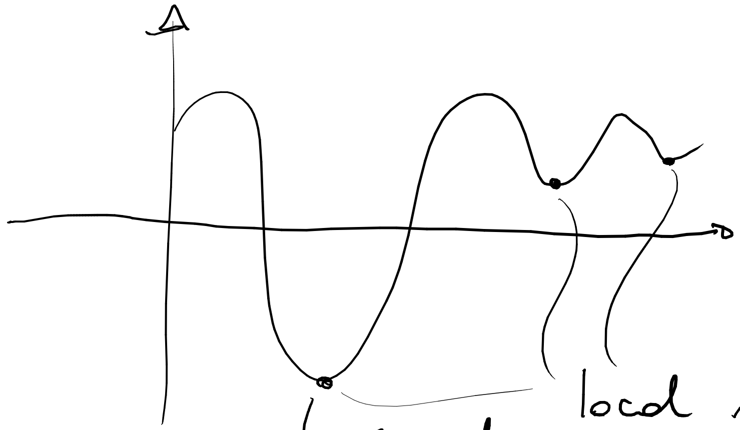
$D \subseteq \mathbb{R}^n$

$|\cdot|$ = Euclidean norm

A point $x^* \in D$ is a LOCAL MINIMUM of f over $D \subseteq \mathbb{R}^n$ if

$\exists\, \varepsilon > 0$ s.t. $\forall x \in D$ satisfying $|x - x^*| < \varepsilon$

$\Rightarrow \boxed{f(x^*) \leq f(x)} \leadsto$ if $f(x^*) < f(x)$ there's a STRICT LOCAL MINIMUM $\forall x \neq x^*$



local minima in their subintervals

$\hookrightarrow$ Global minimum

A point $x^* \in D$ is a GLOBAL MINIMUM of f over $D \subseteq \mathbb{R}^n$ if

for all $x \in D \Rightarrow f(x^*) \leq f(x)$

If a point is either a maximum or a minimum is called EXTREMUM

- **Unconstrained optimization** (no limitations)

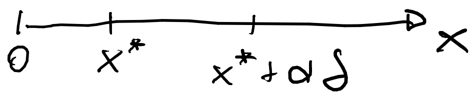
  - first order necessary conditions

All points x sufficiently near x^* in $\mathbb{R}^n$ are in D

Assume $f \in C^1$ and x^* is its local minimum $(f(x^*) \leq f(x))$
Let $\delta \in \mathbb{R}^n$ (arbitrary vector)

Being in the unconstrained case:

$x^* + \alpha\delta \in D \quad \forall \alpha \in \mathbb{R}$ close enough to 0



Let's define a function $k(\alpha) := f(x^* + \alpha\delta)$

So $\boxed{0 \text{ is a minimum of } k}$

First order Taylor expansion of k around $\alpha = 0$

$$k(\alpha) = k(0) + k'(0)\,\alpha + o(\alpha)$$

$$\lim_{\alpha \to 0} \frac{o(\alpha)}{\alpha} = 0$$

$$\Rightarrow \boxed{k'(0) = 0}$$

GOAL

Proof:

I want to prove that demonstrating by contraddiction

assume $k'(0) \neq 0 \Rightarrow \exists\ \varepsilon > 0$ small enough s.t.
for $|\alpha| < \varepsilon$ $\quad |o(\alpha)| < |k'(0)\alpha|$

For these values of α

$$k(\alpha) - k(0) = k'(0)\alpha + o(\alpha) < k'(0)\alpha + |k'(0)\alpha|$$

If we restrict α to have the opposite sign of $k'(0)$

$k(\alpha) - k(0) < 0 \Rightarrow$ Contraddiction $(k'(0) = 0)$

End

$k'(\alpha) = \nabla f(x^* + \alpha\delta)$
$\quad \hookrightarrow \nabla f = (f_{x_1} \dots f_{x_n})^T$ gradient of f

$k'(0) = 0 = \nabla f(x^*) \cdot \delta$
Being δ arbitrary:

$\boxed{\nabla f(x^*) = 0}$ First order necessary condition for optimality

A point x^* satisfying this condition is a stationary point. The result is valid when $f \in C^1$ and x^* is an interior point (inside D).

Therefore when $\alpha = 0$ (not $x^* = 0$), $k(\alpha)$ has a minimum (we are not saying that $f(x^*) = 0$) so $k'(0) = 0$.

- Second order conditions

(Like before but second order)

Assume $f \in C^2$ and x^* its local minimum

Let $\delta \in \mathbb{R}^n$ be an arbitrary vector.

Second order Taylor expansion of k around $\alpha = 0$

$$k(\alpha) = k(0) + k'(0)\alpha + \frac{1}{2} k''(0)\alpha^2 + o(\alpha^2),$$

$$\lim_{\alpha \to 0} \frac{o(\alpha^2)}{\alpha^2} = 0$$

Since $k'(0) = 0 \Rightarrow k''(0) \geq 0$ (not negative)

Proof :

Suppose $k''(0) < 0 \Rightarrow \exists\, \varepsilon > 0$ small enough so that for $|\alpha| < \varepsilon$ $|o(\alpha)^2| < \frac{1}{2}|k''(0)|\alpha^2$

For these values of α, $k(\alpha) - k(0) < 0 \rightarrow$ Contraddiction

We already know that $k'(0) = 0$ from the previous case, therefore $k'' \geq 0$ (non negative)

End

it means that

$$k(\alpha) = f(x^* + \alpha\delta)$$
$$k'(\alpha) = \nabla f(x^* + \alpha\delta)\delta = \sum_{i=1}^{n} f_{x_i}(x^* + \alpha\delta)\delta_i$$

By differentiating both sides of

$$k'(\alpha) = \sum_{i=1}^{n} f_{x_i}(x^* + \alpha \delta)\delta_i \quad \text{with respect to } \alpha$$

$$k''(\alpha) = \sum_{i,j=1}^{n} f_{x_i x_j}(x^* + \alpha\delta)\delta_i \delta_j$$

$$\Rightarrow k''(0) = \sum_{i,j=1}^{n} f_{x_i x_j}(x^*)\delta_i\delta_j = \delta^T \underbrace{\nabla^2 f(x^*)}\delta$$

x^* is a strict local minimum of f

Hessian matrix

$$\nabla^2 f = \begin{pmatrix} f_{x_1 x_1} & \cdots & f_{x_1 x_n} \\ \vdots & \ddots & \vdots \\ f_{x_n x_1} & \cdots & f_{x_n x_n} \end{pmatrix}$$

Second order necessary condition for optimality

$$\boxed{\nabla^2 f(x^*) \geq 0}$$

The result is valid when $f \in C^2$ and x^* is an interior point

Remark:

Being δ arbitrary :

$f \in C^1 \quad \nabla f(x^*) = 0$

$f \in C^2 \quad \nabla^2 f(x^*) \geq 0$

The second order condition distinguishes minima from maxima

- At a local maximum $\nabla^2 f(x^*) \leq 0$ (negative semidefinite)
- At a local minimum $\nabla^2 f(x^*) \geq 0$ (positive semidefinite)

- Weierstrass Theorem (Existence result)

Determining points of global minimum (optimal) of f in D.

$$D = \{ z \in \mathbb{R}^n : h(z) = 0 , g(z) \leq 0 \}$$

means finding every point $z^\circ \in D$ s.t.:

$$f(z^\circ) \leq f(z) , \forall z \in D$$

For the existence of the optimal solutions, Weierstrass theorem gives the sufficient conditions: $C^\circ(D)$

If D is a compact set and f is continuous on D, there exist global minimum (optimal point)

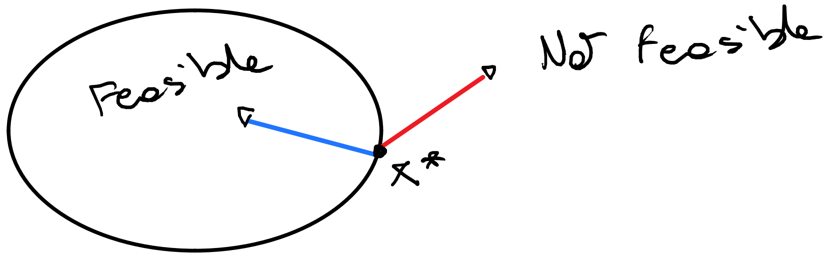
Compact: property of closure and limitation of the domain

- Closure: given by the structure of h and g (they need to be continuous).

- Domain Limitation: it's not guaranteed by the structure of the domain, it depends only on the constraints: it holds only if $\lim_{z \to \infty} h \neq 0$ and/or some components of g are > 0 $\forall z \to \infty$

~ Feasible directions

A vector $\delta \in \mathbb{R}^n$ is a feasible direction at x^* if $x^* + \alpha\delta \in \Delta$ for small enough $\alpha > 0$



$\boxed{\text{If not all directions } \delta \text{ are feasible, then the condition } \nabla f(x^*) = 0 \text{ is no longer necessary for optimality}}$

Previously $\nabla f(x^*) = 0$.
If x^* is a local minimum $\Rightarrow$ $\nabla f(x^*) \cdot \delta \geq 0$
$\forall$ feasible direction δ

~ Procedure for finding a global minimum

1. Find all interior points of Δ satisfying $\nabla f(x^*) = 0$ (stationary points)
2. If f is not differentiable everywhere, include also points where $\nabla f(x^*) = 0$ does not exist (critical points)
3. Find all boundary points satisfying $\nabla f(x^*)\delta \geq 0$ for all feasible δ
4. Compare values at all these candidate points and choose the smallest one.

~ **Convexity**

If f is a convex function and $D \subset \mathbb{R}^n$ is a convex set, a local minimum is automatically a global one and the first order necessary condition (for $f \in C^1$) is also a sufficient condition.

• Constrained optimization

~ Lagrangian

Let $D \subset \mathbb{R}^n$ and $f \in C^1$

$h_1(x) = \dots = h_p(x) = 0$, $h \in C^1$ $\Rightarrow$ Equality constraints
$g_1(x) \leq 0, \dots, g_q(x) \leq 0$, $g \in C^1$ $\Rightarrow$ Inequality constraints

$$\text{Rank}\left\{ \left.\frac{\partial(h, g_a)}{\partial x}\right|_{x^*} \right\} = p + q_a \Rightarrow \text{Regularity condition}$$

the gradients of h and g_a should be linearly independent

g_a = active constraint of g with dimension q_a (the constraint g when it is equal to 0)

Useful to avoid strange conditions like redundant constraints or strange situations

$$\mathcal{L}(x, \lambda_0, \lambda, \eta) = \lambda_0 f(x) + \lambda^T g(x) + \eta^T g(x)$$

↳ the Lagrangian is something like a perturbation

λ: Lagrange multipliers

η: Kuhn-Tucker multiplies

$$\boxed{\text{If } \lambda_0 \neq 0 \text{ (usually } \lambda_0 = 1\text{) the stationary point } x^* \text{ is called } \underline{\text{normal}}}$$

~ First order necessary conditions

$f:\mathbb{R}^n \to \mathbb{R}$ with $f, h, g \in C^1$ with x^* its local minimum

$$\boxed{\left.\frac{\partial \mathcal{L}}{\partial x}\right|_{x^*} = 0^T \qquad \eta_i g_i(x^*) = 0, \ \forall i \qquad \eta_i \geq 0 \ \forall i}$$

If f and g convex and h linear $\Rightarrow$ necessary & sufficient conditions

Proof (only for equality constraints)

Assume that x^* is a local minimum and a regular point
$\to \nabla h_i, \ i = 1 \ldots m$ are linearly independent in x^*

Let $x(\alpha) \in D$ such that $x(0) = x^*$ with $x(\alpha)$ a family of curves passing through x^* $(\alpha \in \mathbb{R})$

Consider the function

$k(\alpha) = f(x(\alpha))$, $k(0) = f(x^*)$ is a minimum of k

$k'(\alpha) = \nabla f(x(\alpha)) x'(\alpha)$, $k'(0) = \nabla f(x^*) x'(0) = \nabla f(x^*) \delta = 0$ $\forall$ tangent vector δ

Consider the Taylor expansion (1-st order)

$x(\alpha) = \underbrace{x'(0)}_{\text{tangent vector to } D \text{ at } x^*} \alpha + o(\alpha)$

The tangent space to D at x^* is characterized by

$h_i(x(\alpha)) = 0 \quad \forall \alpha$ with $i = 1, \ldots, m$

$$\left.\frac{d}{d\alpha}\right|_{\alpha=0} h_i(x(\alpha)) = \nabla h_i(x^*) x'(0) = \underbrace{\nabla h_i(x^*) \delta = 0}_{} \quad \forall \delta \in \mathbb{R}^n$$

The tangent vectors to D at x^* are exactly δ for which the condition holds

$$\Rightarrow \boxed{\nabla f(x^*) \in \operatorname{span}\{\nabla h_i(x^*), \ i = 1 \ldots m\}}$$

There exist real numbers $\lambda_1^*, \ldots, \lambda_n^*$ such that:

$$\nabla f(x^*) + \lambda_1^* \nabla h_1(x^*) + \ldots + \lambda_n^* \nabla h_n(x^*) = 0$$

End

~ Second order sufficient condition (not always applicable)

Let $x^* \in D$ and $f, h, g \in C^2$ and assume the conditions

↳ very regular

$$\left.\frac{\partial \mathcal{L}}{\partial x}\right|_{x^*} = 0^T \qquad \eta_i g_i(x^*) = 0, \quad \eta_i \geq 0 \quad \forall i$$

x^* is a strict local minimum if

$$\delta^T \left.\frac{\partial^2 \mathcal{L}}{\partial x^2}\right|_{x^*} \delta > 0 \quad \forall \delta \text{ such that } \left.\frac{d h_i(x)}{dx}\right|_{x^*} \delta = 0 \quad i = 1 \dots p$$

- Bordered Hessian (Not used in this course)

$$H = \begin{pmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial h_j}{\partial x_i} \\ \left(\frac{\partial h_j}{\partial x_i}\right)^T & 0 \end{pmatrix}$$

↳ borders

A point x^* in which $\nabla \mathcal{L} = 0$ and $\det(H) \neq 0$ is called a non-degenerate critical point of the constrained problem

- **Function spaces**

Functional (function of functions): $J: V \to R$ with $A \subseteq V$
V: vector space

~ **Norm $\|\cdot\|$**

It is a real valued function on V

- Positive definite $\|y\| > 0$ if $y \neq 0$
- Homogeneous $\|\lambda y\| = |\lambda| \cdot \|y\|$ $\forall \lambda \in R, y \in V$
- Satisfies the triangle inequality $\|y+z\| \leq \|y\| + \|z\|$

~ **Distance or metric**

$$d(y,z) = \|y-z\|$$

~ **Strong extreme (0-Norm)** → A strong minimum is also a weak one (not the inverse)

On the space $C^0([a,b], R^n)$

$$\|y\|_0 = \max_{a \leq x \leq b} |y(x)| \Rightarrow \text{Euclidean norm}$$

$$\boxed{\text{Extreme of } J \text{ wrt the 0-Norm are strong extreme}}$$

~ **Weak extreme (1-Norm)**

On the space $C^1([a,b], R^n)$

$$\|y\|_1 = \max_{a \leq x \leq b} |y(x)| + \max_{a \leq x \leq b} |y'(x)|$$

$$\boxed{\text{Extreme of } J \text{ wrt the 1-Norm are weak extreme}}$$

~ **k-th Norm**

On the space $C^\ell([a,b], R^n)$, $\ell \geq k \Rightarrow \|y\|_p = \left(\int_a^b |y(x)|^p dx\right)^{1/p}$

~ Definition of extreme

$z^* \in A$ is a local minimum of J over A if
$\exists \varepsilon > 0 : \forall z \in A$ such that $\|z - z^*\| < \varepsilon \Rightarrow J(z^*) \leq J(z)$

~ Variation

Notation: δ = variation (like the incremental ratio)
$\downarrow$
derivative

Let $z + \alpha\eta$, $\eta \in V$, $\alpha \in \mathbb{R}$ a function in V.

It is an admissible perturbation wrt a subset A
if $z + \alpha\eta \in A \quad \forall \alpha \to 0$

The first variation of J at z is the linear function
$\delta J|_z : V \to \mathbb{R}$ such that $J(z+\alpha\eta) = J(z) + \delta J|_z(\eta)\alpha + o(\alpha)$

or defined as:

$$\boxed{\delta J|_z(\eta) = \lim_{\alpha \to 0} \frac{J(z+\alpha\eta) - J(z)}{\alpha}}$$

- First order necessary condition for optimality

$$\boxed{\delta J|_{z^*}(\eta) = 0}$$

- Second order necessary condition

$$\boxed{\delta^2 J|_{z^*}(\eta) \geq 0}$$

with a second variation of J at z if
$J(z+\alpha\eta) = J(z) + \delta J|_z(\eta)\alpha + \delta^2 J|_z(\eta)\alpha^2 + o(\alpha^2)$

~ Weierstrass theorem

$\boxed{\text{If } A \text{ is a compact set and } J \text{ is continuous on } A\text{, there exist global minimum}}$

~ Convexity

$A \subset V$ convex with $J : A \to \mathbb{R}$

J is a convex functional on A if

$$J(\alpha z' + (1-\alpha) z'') \leq \underbrace{\alpha J(z') + (1-\alpha) J(z'')}_{\text{z' and z'' other points}}$$

linear combination

$\boxed{\text{If } J \text{ is a convex function and } A \subset V \text{ is a convex set, a local minimum is automatically a global one and the first order conditions are necessary and sufficient conditions for a minimum}}$

increment

If $A = \mathbb{R}^n$ and $J \in C^1(A)$: convexity $\Leftrightarrow J(z) \geq J(z') + \left.\frac{dJ}{dz}\right|_{z'} (z - z')$

If $A = \mathbb{R}^n$ and $J \in C^2(A)$: convexity $\Leftrightarrow (z - z')^T \left.\frac{d^2 J}{dz^2}\right|_{z'} (z - z') \geq 0$