Optimization and Optimal Control

Engineers like to do things well

$\Rightarrow$ $\underbrace{\text{optimization.}}$

Contents of this lecture:

- Unconstrained optimization, including algorithms:
  - steepest descent,
  - conjugate gradient,
  - Newton, pseudo-Newton
- Solution of nonlinear vector equations
- Generalized matrix inverses and least-squared error solution of $y = Ax$
- Duality and constrained optimization, (convex analysis used)
- Introduction to interior-point methods
- Introduction to dynamical optimization → optimal control theory

Some notation:

Suppose:

a statement α $\Rightarrow$ $\square$ — a contradiction

(eg. $(\beta=0)\wedge(\beta>0)$)

Conclusion: α must be false

Hence one can prove that θ is true by showing

$\neg\theta \Rightarrow \square$ — a contradiction

not

called a proof by contradiction

A motivation for a basic optimization problem

$$x \longrightarrow \boxed{\text{System or design}} \longrightarrow V(x)$$

x = vector of parameters

$V(x)$ = system's operating cost or cost of design

Could be a usual financial cost perhaps including penalties for not meeting specifications, etc.

A basic optimization problem

parameter vector $x \in \mathbb{R}^n$

cost function $V : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\boxed{\begin{array}{l} \text{minimize } V(x) \\ \; x \in F \end{array}}$$

feasible set = the set of x's which satisfy the constraints

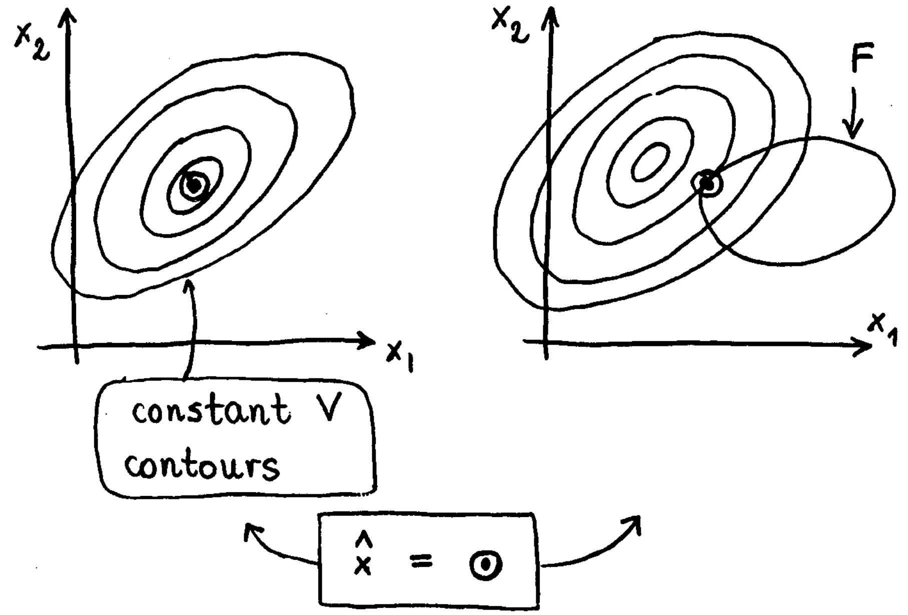
= eg. the set of allowable parameter vectors, bearing in mind constraints

e.g. temperature constraints for chemical processes, size and strength constraints for mechanical engineering designs, etc.

Examples : $n = 2$

unconstrained opt. $F = \mathbb{R}^2$

constrained opt. $F \subsetneq \mathbb{R}^2$



Meaning of optimization:

(1.1) $\hat{x}$ = a global minimizer of V on F

iff

$$[\hat{x} \in F] \wedge [V(x) \geq V(\hat{x}),\ \forall x \in F]$$

Set of all global minimizers :

$$\boxed{\underset{x \in F}{\arg\min}\ V(x)}$$

Optimal value :

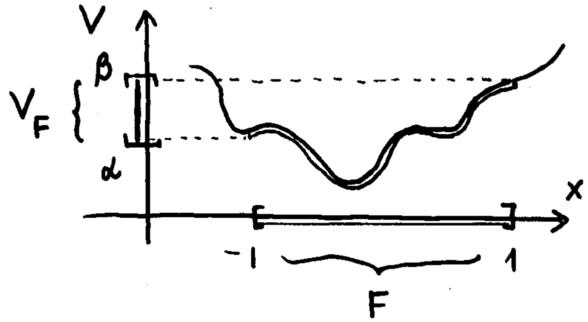
$$\boxed{\hat{V} \triangleq \min_{x \in F} V(x)}$$

Existence of an $\hat{x}$?

When might there not be a g.m. ?

$$\boxed{\begin{aligned}(\alpha, \beta] &\triangleq \{z \in \mathbb{R} : \alpha < z \leq \beta\} \\ [\alpha, \beta) &\triangleq \{z \in \mathbb{R} : \alpha \leq z < \beta\}\end{aligned}}$$

Cost value set: $V_F \triangleq \{V(x) : x \in F\}$



(1.2) **Th.** (i) $\exists$ g.m. if eg. $V_F = [\alpha, \beta]$

(ii) $\nexists$ g.m. if eg. $V_F = (\alpha, \beta]$

<u>Proof</u> [of (ii)] : $\boxed{V_F = (\alpha, \beta]}$

We will show that :

$\# : [\hat{x} = \text{a g.m.}] \Rightarrow \square$ a contradiction

therefore : $\nexists$ any g.m. because $\hat{x}$ is just a name for any g.m.

Proof of # :

Suppose $[\hat{x} = \text{a g.m.}]$

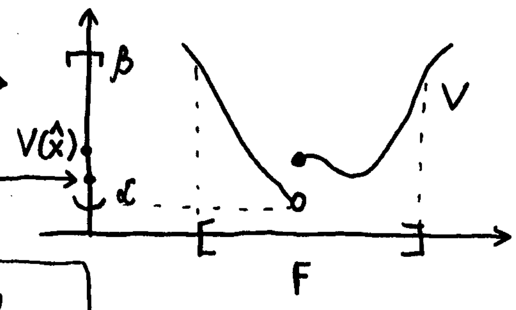
$\Downarrow$

$[\hat{x} = \text{a g.m.}] \wedge [V(\hat{x}) \in V_F]$

$\Downarrow$

$\Downarrow$

We have this situation →



$\exists$ some θ — because $V(\hat{x}) > \alpha$ as $V(\hat{x}) \in V_F = (\alpha, \beta]$

$\Downarrow$

$[\hat{x} = \text{a g.m.}] \wedge [\exists\, \theta \in V_F \text{ with } \theta < V(\hat{x})]$

$\Downarrow$ because $V_F = \{V(x) : x \in F\}$

$[\hat{x} = \text{a g.m.}] \wedge [\exists\, \tilde{x} \in F \text{ with } V(\tilde{x}) = \theta < V(\hat{x})]$

$\Downarrow$ because $\tilde{x}$ is feasible and gives lower V than $\hat{x}$ does, which contradicts the optimality of $\hat{x}$

$\square$ a contradiction

QED

Bounded sets

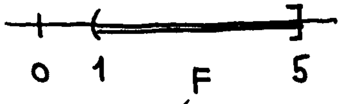
(1.3) $F \subset \mathbb{R}^n$ called bounded iff :

$\exists\ r \in [0,\infty)$ so that

$\|x\| < r \ , \ \forall x \in F$

e.g. $F \subset \mathbb{R}$

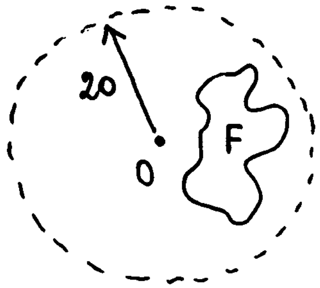
$\boxed{\begin{array}{l} \|x\| = |x| \\ \text{as } x = \text{scalar} \end{array}}$



$\boxed{\begin{array}{l} \text{bounded} \\ \text{e.g. use } r = 6 \end{array}}$

e.g. $F \subset \mathbb{R}^2$

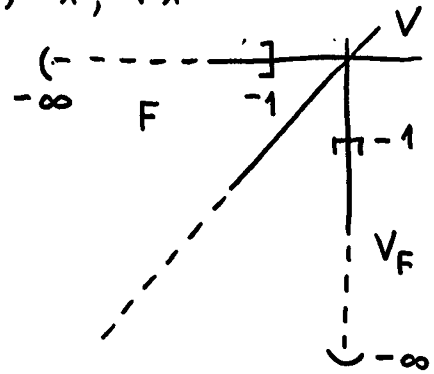
$\boxed{\begin{array}{l} \text{bounded} \\ \text{e.g. use } r = 20 \end{array}}$



I Fact : $\boxed{\begin{array}{l} \text{a g.m. } \hat{x} \text{ might} \\ \text{not exist if } F \text{ is} \\ \text{NOT bounded} \end{array}}$

e.g. $F \subset \mathbb{R}$; $V(x) = x, \ \forall x$

$F = (-\infty, -1]$



$V_F = (-\infty, -1]$

$\Downarrow$

$\nexists$ a g.m.

$\boxed{\begin{array}{l} \text{because} \\ F \text{ is not} \\ \text{bounded} \end{array}}$

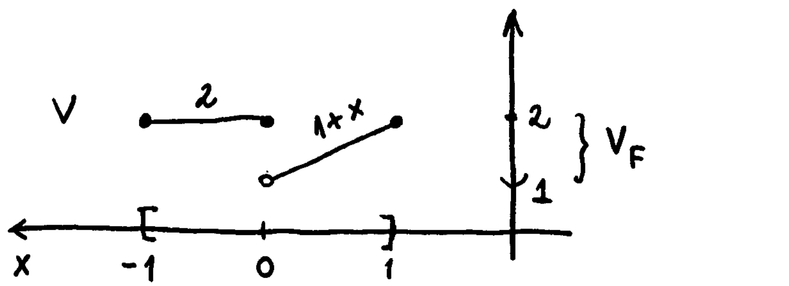
$\boxed{\begin{array}{l} \text{by Th. (1.2)} \\ \text{because } V_F = (\alpha, \beta] \\ \text{for } \alpha = -\infty, \beta = -1 \end{array}}$

$\boxed{\begin{array}{l} -\infty \text{ has occured here} \\ \text{because } F \text{ is not} \\ \text{bounded} \end{array}}$

II Fact: a g.m. $\hat{x}$ might not exist if $V \neq \underbrace{\text{continuous on } F}_{\boxed{V \text{ is continuous at every } x \in F}}$

e.g. $F = [-1, 1] \subset \mathbb{R}$

$$V(x) = \begin{cases} 2, & x \le 0 \\ 1+x, & x > 0 \end{cases}$$



$V_F = (1, 2]$

$\Downarrow$ — Th. (1.2)

no g.m. — because of the discontinuity of V

Limit points and closed sets

(1.4) $p \in \mathbb{R}^n$ is called a Lt. pt. of $F \subset \mathbb{R}^n$

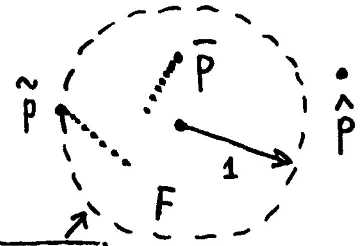
iff: there exist points $q \in F$, $q \neq p$ which are arbitrarily near p #

i.e. iff: you can get as close as you like to p while remaining in F (and while not being equal to p)

e.g. $F = \{x \in \mathbb{R}^2 : \|x\| < 1\}$

	$\in F$	Lt. pt.
$\bar{p}$	Y	Y
$\tilde{p}$	N	Y
$\hat{p}$	N	N

not closed



F is called closed if it contains all its limit points

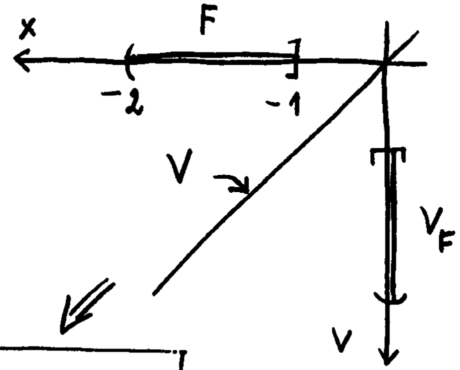
(but $F = \{x \in \mathbb{R}^2 : \|x\| \le 1\}$ <u>is</u> closed

more precisely (for math. friends only):
iff: for every $\varepsilon > 0$,
$\exists\, q \in F$, with $q \neq p$, such that $\|p - q\| < \varepsilon$

III Fact: there might not exist a g.m. $\hat{x}$ if $F \neq$ closed

e.g. $F = (-2, -1]$, $V(x) = x$, $\forall x$

not closed:
$-2 =$ lt. pt.
$-2 \notin F$



no g.m.
because V_F has the form $(\alpha, \beta]$, which has happened because $F \neq$ closed

Examples shown:

might not be a g.m. $\hat{x}$ if:

- $F \neq$ bounded
- $F \neq$ closed
- $V \neq$ continuous

Suggests:

(1.5) **Th**: $\exists$ a g.m. $\hat{x}$ if:

- $F =$ closed & bounded
- $V =$ continuous on F

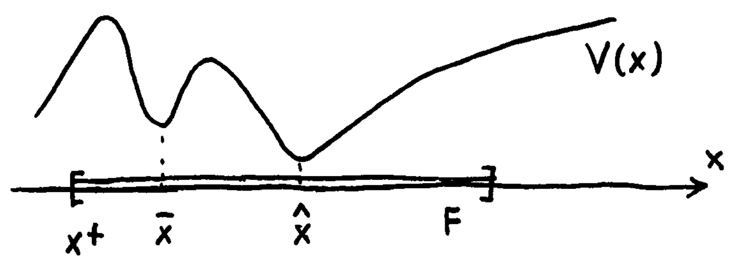
Proof: too hard

But: Existence of a g.m. matters as there is not any hope of using an algorithm to find a g.m. if one does not exist! Also, this topic deepens understanding of what optimization is about.

Assumption:

For V & F in this course: $\exists$ a g.m. $\hat{x}$

Local minimizers



$\hat{x}$ = the g.m.

$\hat{x}, x^+, \bar{x}$ = local minimizers

(1.6) $\tilde{x}$ = a local minimizer of V on F

iff :

$$[\tilde{x} \in F] \wedge \left[V(x) \geq V(\tilde{x}), \ \forall \ \underline{\text{feasible } x} \ \underline{\underline{\text{near } \tilde{x}}} \right]$$

i.e. x such that

$[x \in F] \wedge [\| x - \tilde{x} \| < \varepsilon]$

for suitably small $\varepsilon > 0$

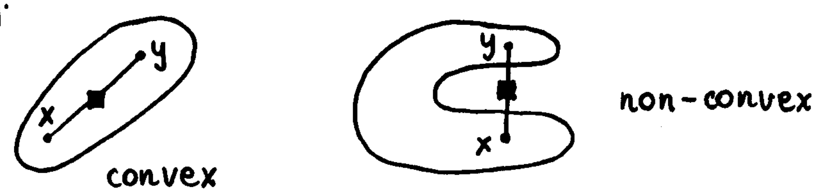
Convexity

... a useful property of many V & F.

A set F is convex iff

every point along the line joining any two points in F is itself in F

e.g.



Algebraic characterization :

(1.7) F is convex iff

$$\underbrace{\alpha x + (1-\alpha) y}_{\blacksquare} \in F$$

$\forall \ \alpha \in [0,1]$

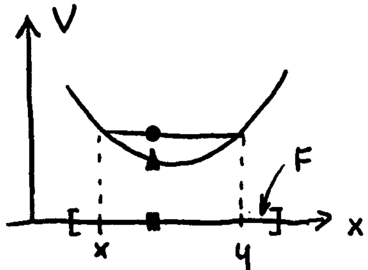
$\forall \ x, y \in F$

Convex functions

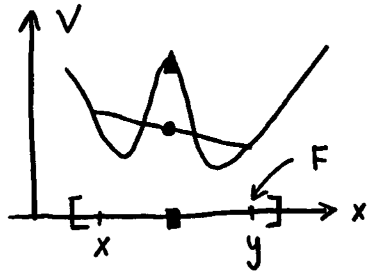
(1.8) A function $V : \mathbb{R}^n \to \mathbb{R}$ is called <u>convex on a convex set F</u> iff :

$$\underset{\blacktriangle}{V}[\underbrace{\alpha x + (1-\alpha) y}_{\blacksquare}] \le \underbrace{\alpha V(x) + (1-\alpha) V(y)}_{\bullet}$$

$$\forall \alpha \in [0,1], \quad \forall x, y \in F$$



V = convex on F



$V \neq$ convex on F

$\Downarrow$

$$\boxed{\text{So: for } V \text{ to be convex on } F: V \text{ must lie below (i.e. } \le \text{) the straight line interpolating between } V(x) \text{ and } V(y) \text{ everywhere between } x \text{ and } y, \text{ and for all } x, y \in F.}$$

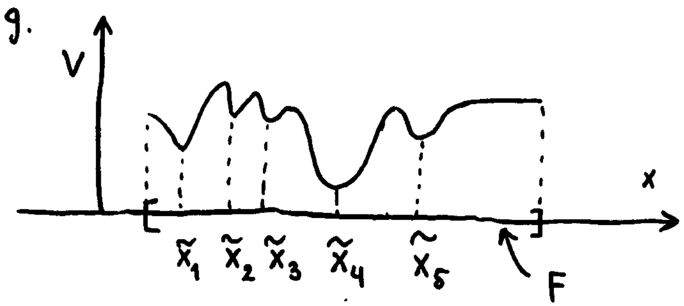
Significance of convexity

Most opt. algs. find only a local minimizer (l.m.)

In general : to find a g.m. :

- first find all l.m.s — generally impossible
- select the best

e.g.



$\tilde{x}_1$ $\tilde{x}_2$ $\tilde{x}_3$ $\tilde{x}_4$ $\tilde{x}_5$

$\hat{x} = \tilde{x}_4$

So, in general

finding a g.m. $\neq$ possible, BUT

$\Downarrow$

(1.9) **Th** Suppose :

V is convex on convex F

$\exists$ a g.m.

Then : all l.m.s of V on F are also g.m.s.

SO :

$\boxed{\text{For convex } V \;\&\; F\text{: we can stop optimizing once any l.m. is found, as it must also be a g.m.}}$

Proof : (by showing that any l.m. $\bar{x}$ is actually a g.m.)

(only for $n = 1$, but generalizes easily)

Say : $\hat{x}$ = a g.m.

$\bar{x}$ = a l.m. with $\hat{x} \neq \bar{x}$

Clearly : $V(\bar{x}) \geqslant V(\hat{x})$

but

(#) $[V(\bar{x}) > V(\hat{x})] \Rightarrow \square$ a contra-diction

hence $V(\bar{x}) = V(\hat{x})$

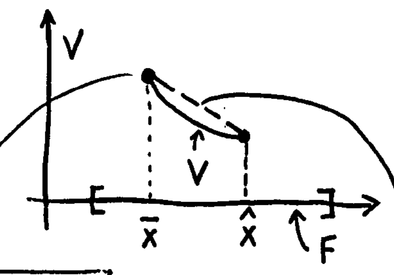
$\Rightarrow$ l.m. $\bar{x}$ has the same cost as g.m. $\hat{x}$

$\Rightarrow \bar{x}$ = a g.m. too

QED.

Proof of (#) : $[V(\bar{x}) > V(\hat{x})]$

$\Downarrow$



$\Rightarrow \square$ a contra-diction

contradicts
$\bar{x}$ = a l.m.

this situation arises because, as V is a convex function on F, V lies below the line interpolating between $V(\bar{x})$ and $V(\hat{x})$.

QED