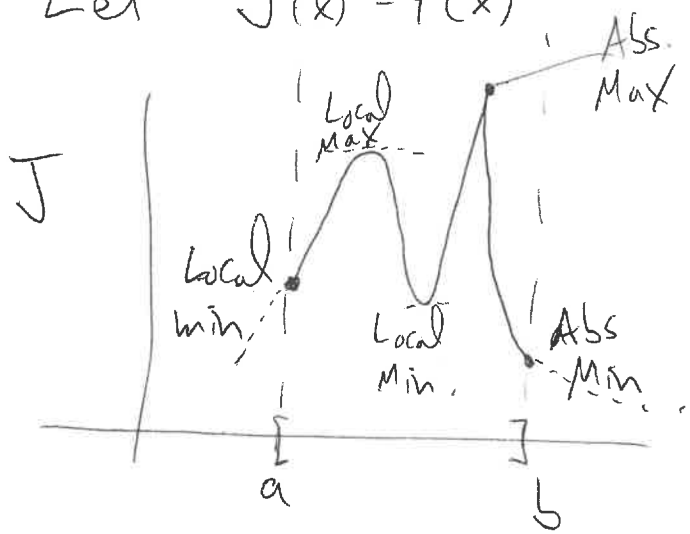
Parametric Optimization (Chapter 1 in Lawden)

Necessary & Sufficient Conditions for a cost function to have Maximum or Minimum.

No Constraints.

Consider a function $f(x)$, continuous over an $[a, b]$

Let $J(x) = f(x)$



The abs. max/min is either at the endpoints, or in the interior. In the interior, either has 0 slope or discontinuous slope.

Consider a scalar function of n indep. variables

$$\bar{x} = (x_1, x_2, \cdots, x_n)$$

$$J(\bar{x}) = f(x_1, x_2, \cdots, x_n) = f(\bar{x}) = f(x_i;\ i=1,\cdots,n)$$

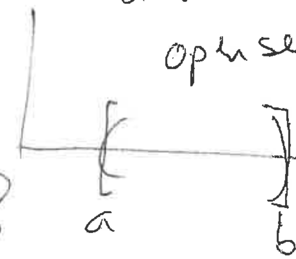
$J(\bar{x})$ is defined in a region R of x_i

The location of a particular point $\bar{x} \in R$ is considered to be a valid point.

A point $\bar{x} \in R$ where J acheives a max/min (extremum) is determined by the following:

<u>Thm</u>: A <u>continuous</u> function $f(\bar{x})$ of n <u>independent</u> variables x_i attains a max/min in the <u>interior</u> (open set) of a region R <u>only</u> at those values of x_i where the partials of f wrt the x_i, $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \cdots$ either <u>vanish</u> simultaneously, or at which one or more of these are discontinuous.

Thm: If the set R is compact, the extremum of $f(\bar{x})$ either occurs in the interior of R or at the boundaries.

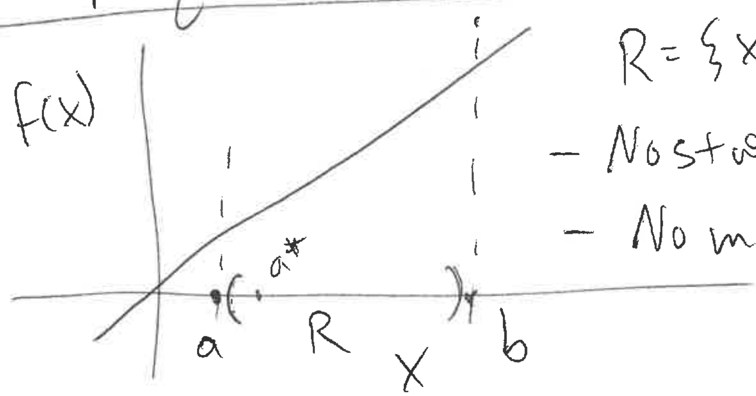


compact − open set $= \{a, b\}$

Weierstrass Existence Conditions

If $f(\bar{x})$ is continuous over $\bar{x} \in R$, and R is compact, then $f(\bar{x})$ has a max/min value over the domain R

Examples of functions w/o max or min



$R = \{x \mid a < x < b\}$

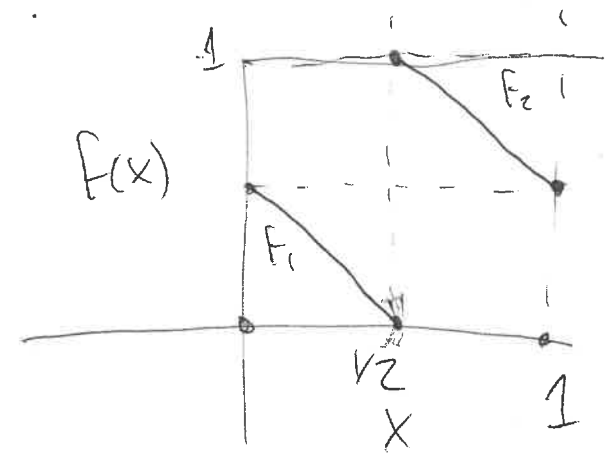
- No stationary pts in interior
- No max or min

If I say $a^* = \epsilon$

$f(a^*) = \epsilon$

$\frac{a^*}{2}$

Continuous?



$x \in [0, 1]$

$f(x = \frac{1}{2}) \equiv$ Can't have 2 values...

Either

(A) $f(x) = \begin{cases} f_1(x) & x \in [0, \frac{1}{2}) \\ f_2(x) & x \in [\frac{1}{2}, 1] \end{cases}$

(B) $f(x) = \begin{cases} f_1(x) & x \in [0, \frac{1}{2}] \\ f_2(x) & x \in (\frac{1}{2}, 1] \end{cases}$

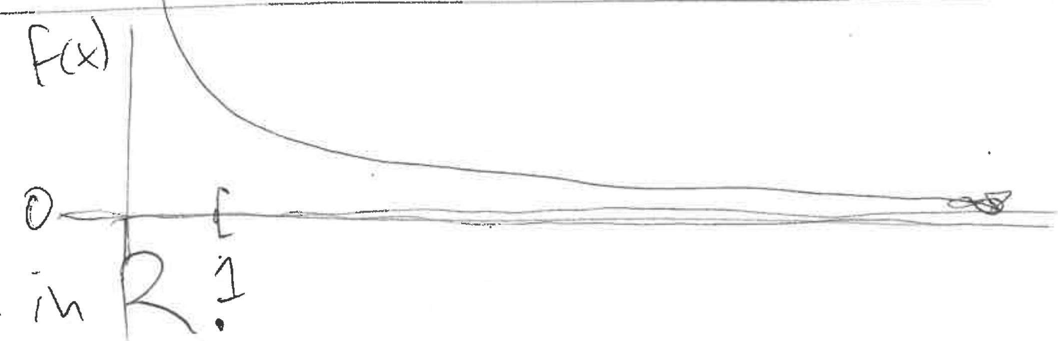
The discontinuity gives us an "interior open set" that messes things up....

Option (A) = then $f(x)$ has no minimum value, has max of 1

(B) $\Rightarrow$ " " " " maximum ", " min of 0

Even extends to the real line

$f(x) = \frac{1}{x}$ $R = [1, \infty)$



Does not have a minimum in R!

Infimum of a set S: Greatest element of an ordered set less than "S"
Supremum of a set S: Least " " " " "greater" "

Inf $f(x) = 0$. How to prove?
If $\lim_{x \to \infty} f(x) = f_\infty$ exists, can perform "one-point compactification"
which allows us to define a closed version of

$$R = [0, \infty], \quad f(x) = \begin{cases} 1/x & x \in [1, \infty) \\ f_\infty & x = \infty \end{cases}$$

Necessary & Sufficient Conditions for optimality

$$J = f(\vec{x}) \; ; \; \vec{x} \in R \subset \mathbb{R}^n$$

Parametric Optimization : Nec. & Suff. Conditions

Given $J = f(\vec{x})$ $\quad \vec{x} \in R \subset \mathbb{R}^n$

Problem Statement:

- If $f(\vec{x})$ is continuous (C^0) and R is compact,

Find $\boxed{J^* = \min_{\vec{x} \in R}/\max f(\vec{x}) \quad \& \quad \vec{x}^* = \arg \binom{\min}{\max}_{\vec{x} \in R} f(\vec{x})}$

- If $f(\vec{x})$ is not continuous or R is not compact

Find $J^* = \inf_{\vec{x} \in R}/\sup f(\vec{x})$, $\quad \vec{x}^* = \arg \frac{\inf}{\sup}_{\vec{x} \in R} f(\vec{x})$, but

$\vec{x}^*$ is not nec. defined or $\vec{x}^* \notin R$

Nec. & Suff. Conditions : Global, Local Smooth, Local Non-Smooth

Use "minimum" for definitiness . . . . .

Global

Nec. Cond: IF $\vec{x}^*$ is a minimizer of $J = f(\vec{x})$,

Then $J(\vec{x}^* + \Delta\vec{x}) \geq J(\vec{x}^*)$, $\forall\ \vec{x}^* + \Delta\vec{x} \in R$

IF $\vec{x}^*$ is the unique global minimizer,

then $J(\vec{x}^* + \Delta\vec{x}) > J(\vec{x}^*)$ $\forall\ \vec{x}^* + \Delta\vec{x} \in R$

$|\Delta\vec{x}| > 0$

- Suff. Condition

$\vec{x}^*$ is a minimizer of J if

$$J(\vec{x}^* + \Delta\vec{x}) \geq J(\vec{x}^*) \quad \forall\ \vec{x}^* + \Delta\vec{x} \in R$$

$\vec{x}^*$ is the unique minimizer if

$$J(\vec{x}^* + \Delta\vec{x}) > J(\vec{x}^*) \quad \forall\ \vec{x}^* + \Delta\vec{x} \in R$$

$$|\Delta\vec{x}| > 0$$

Difficult to Solve

Exception is if $f(\vec{x})$ is convex.

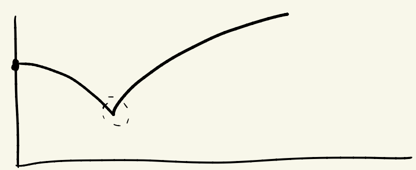
Local Conditions: Can be checked explicitly ...

Assume our candidate points $\vec{x} \subset R$, but not $\vec{x} \not\subset \partial R$

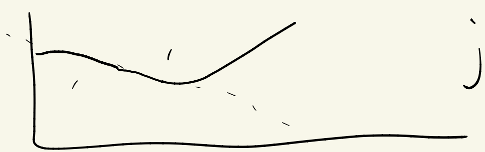
Smooth $f(\vec{x})$: A function $f(\vec{x})$ is "C^n" if its first n- derivatives are continuous in its domain.

A function is smooth is $n > 1$ (2)

C^0



C^1



; $C^1 = \frac{\partial f}{\partial \vec{x}}$ is continuous

$C^2 = \frac{\partial^2 f}{\partial \vec{x}^2}$ " "

Let $\Delta\vec{x} = \delta\vec{x}$; $|\delta\vec{x}| << 1$

Consider functional variations in the <u>nbhd</u> of a candidate point, $\vec{x}^*$.

Nec. If $\vec{x}^*$ is a local minimizer of J, then

$$\underbrace{J(\vec{x}^* + \delta\vec{x})}_{\text{Taylor Series}} \geq J(\vec{x}^*) \quad \forall\ \delta\vec{x} \ni |\delta\vec{x}| << 1$$

$$\cancel{J(\vec{x}^*)} + \delta J(\vec{x}^*) + \delta^2 J(\vec{x}^*) + \cdots \geq \cancel{J(\vec{x}^*)}$$

$$\boxed{\delta J(\vec{x}^*) + \delta^2 J(\vec{x}^*) + \cdots \geq 0}$$

$$\delta J^* = \left.\frac{\delta F}{\delta \vec{x}}\right|_{\vec{x}^*} \cdot \delta\vec{x}$$

$$\delta^2 J^* = \frac{1}{2!}\, \delta\vec{x}^T \left.\frac{\delta^2 F}{\delta \vec{x}^2}\right|_{\vec{x}^*} \delta\vec{x}$$

$$\delta^2 J^* = \frac{1}{2!}\, \delta\vec{x} \cdot \left.\frac{\delta^2 F}{\delta \vec{x}^2}\right|_{\vec{x}^*} \cdot \delta\vec{x}$$

Consider $O(\delta \vec{x})$ . . .

$$\delta^* J = \frac{\partial f}{\partial \vec{x}}\bigg|_{\vec{x}^*} \cdot \delta \vec{x} \geq 0 \quad 0 < |\delta \vec{x}| << 1$$

If $\left| \frac{\partial f}{\partial \vec{x}}\bigg|_{\vec{x}^*} \right| \neq 0$, there always exists a $\delta \vec{x}$, such that

$$\delta^* J < 0 \; ; \; \delta \vec{x}' = -\frac{\partial f}{\partial \vec{x}}\bigg|_{\vec{x}_*} \epsilon$$

Thus a Nec. Condition for $\vec{x}^*$ to be a minimizer is

$$\boxed{(A) \quad \frac{\partial f}{\partial \vec{x}}\bigg|_{\vec{x}^*} = \vec{0}}$$

Can't tell if its a min, max, saddle

Condition becomes $\boxed{\delta^2 J^* \geq 0}$ or or $\boxed{\frac{1}{2!}\,\delta\vec{x}\cdot\left.\frac{\partial^2 F}{\partial \vec{x}^2}\right|_{*}\cdot\delta\vec{x} \geq 0}$

Depends on the properties of the matrix $\left.\left(\frac{\partial^2 F}{\partial \vec{x}^2}\right)\right|_{\vec{x}^*} = F^*$

F^* is symmetric, A symmetric matrix F^* is positive definite

if $\delta x^T F^* \delta x > 0 \quad \forall\ |\delta\vec{x}| \neq 0$

$\Rightarrow$ Implies that all eigenvalues of F^* are positive

---

F^* is semi-definite if $\delta x^T F^* \delta x \geq 0 \quad \forall\ |\delta\vec{x}| \neq 0$

$\Rightarrow$ Implies is that all eigenvalues are ≥ 0, some zero eigenvalues

F^* is indefinite if $\delta x^T F^* \delta x \gtrless 0 \Rightarrow$ eigenvalues are pos & neg $\rightarrow$ Saddle point

B 1: If $\vec{x}^*$ is a candidate local minimizer then F^* is positive semi-definite, (need to check higher orders)

2: If $\vec{x}^*$ is a unique local minimizer, then F^* is positive definite

---

Suff. Cond.

$\vec{x}^*$ satisfies A, and B2

Global Conditions: If $f(\vec{x})$ is a convex function & $\vec{x}^*$ satisfies either N+S condition, $\vec{x}^*$ is a global minimizer.

A function $f(\vec{x})$ is convex if

$$f\left(t \cdot \vec{x}_1 + (1-t)\,\vec{x}_2\right) \leq t \cdot f(\vec{x}_1) + (1-t) \cdot f(\vec{x}_2)$$

$$\forall\, \vec{x}_1, \vec{x}_2 \subset R \qquad \left(t\,\vec{x}_1 + (1-t)\,\vec{x}_2 \in R \right)$$

If $\left.\dfrac{\partial^2 f}{\partial \vec{x}^2}\right|_{\vec{x}}$ is positive definite $\forall\, \vec{x} \in R$, then f is convex

$$\left.\frac{\partial^2 f}{\partial \vec{x}^2}\right|_{\vec{x}} > 0$$

For non-smooth functions or boundary points we define

$$\boxed{\underline{\text{Gateaux Derivative}} \quad f(\vec{x}) \qquad df_+(\vec{x}_0, \vec{u}) = \lim_{\substack{\lambda \to 0^+ \\ \lambda > 0}} \frac{1}{\lambda}\left[f(\vec{x}_0 + \lambda \vec{u}) - f(\vec{x}_0) \right]}$$

$\vec{u} \equiv$ "search direction"

Can describe the neighborhood of $f(\vec{x}_0)$ as

$$\boxed{f(\vec{x}) = f(\vec{x}_0) + \lambda\, df_+(\vec{x}_0, \vec{u}) + O(\lambda^2)}$$

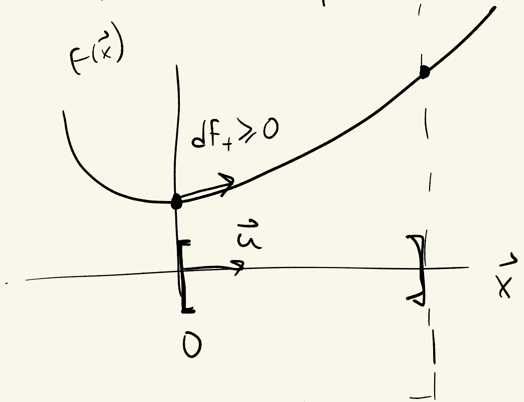
If $f(\vec{x})$ is C^1, then $df_+(\vec{x}_0, \vec{u}) = \left.\frac{\partial f}{\partial \vec{x}}\right|_{\vec{x}_0} \cdot \vec{u} \Rightarrow$

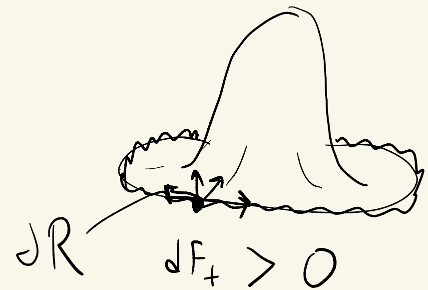
Boundary Points

If $\vec{x}^* \in \partial R$, it is a local minima

if $df_+(\vec{x}^*, \vec{u}) \geq 0$, $\forall$ $\vec{x}^* + \epsilon \vec{u} \in R$

$\epsilon > 0$



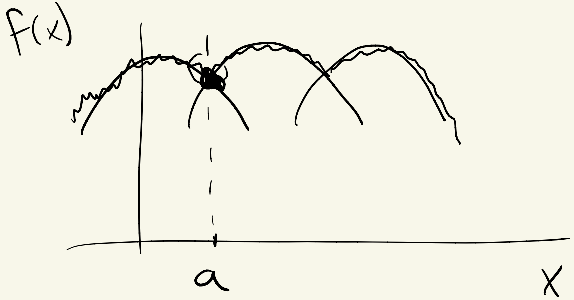


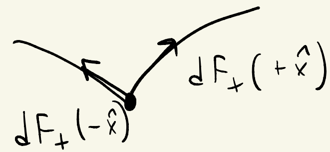
Non-smooth Function:

N.C. If $\vec{x}^* \in R$ is a local minimizer of $f(\vec{x})$, then

$$df_+(\vec{x}^*, \vec{u}) \geq 0 \quad \forall \vec{u} \in \mathbb{R}^n$$

$df_+(\vec{x}^*, \vec{u}) > 0$ " " " " its a unique minimizer





Solve Parametric Optimization Problems

– Nec. Condition Solver

IF $f(\vec{x})$ is smooth, solve $\boxed{\left.\frac{\partial F}{\partial \vec{x}}\right|_{\vec{x}^*} = \vec{0}}$

Use Newton-iteration method.

Given $\vec{x}_0$, where $\left.\frac{\partial F}{\partial \vec{x}}\right|_{\vec{x}_0} \neq \vec{0}$; find $\delta \vec{x}$ such that $\boxed{\left.\frac{\partial F}{\partial \vec{x}}\right|_{\vec{x}_0 + \delta \vec{x}} = \vec{0}}$

Assume $|\delta \vec{x}| \ll 1$

$$\left.\frac{\partial f}{\partial \vec{x}}\right|_{\vec{x}_0+\delta\vec{x}} = \underbrace{\left.\frac{\partial f}{\partial \vec{x}}\right|_{\vec{x}_0} + \left.\frac{\partial^2 f}{\partial \vec{x}^2}\right|_0 \cdot \delta\vec{x}}_{= \vec{0}} + \cdots\cdots = \vec{0}$$

$$\delta\vec{x} = -\left[\frac{\partial^2 f}{\partial \vec{x}^2}\right]\Bigg|_{\vec{x}_0}^{-1} \cdot \left(\frac{\partial f}{\partial \vec{x}}\right)\Bigg|_{\vec{x}_0} \Rightarrow \text{First order correction.}$$

For $\vec{x}_1 = \vec{x}_0 + \delta\vec{x} \Rightarrow$ check if $\left|\left.\frac{\partial f}{\partial \vec{x}}\right|_{\vec{x}_1}\right| < \epsilon \simeq$ accuracy tolerance

If not, repeat

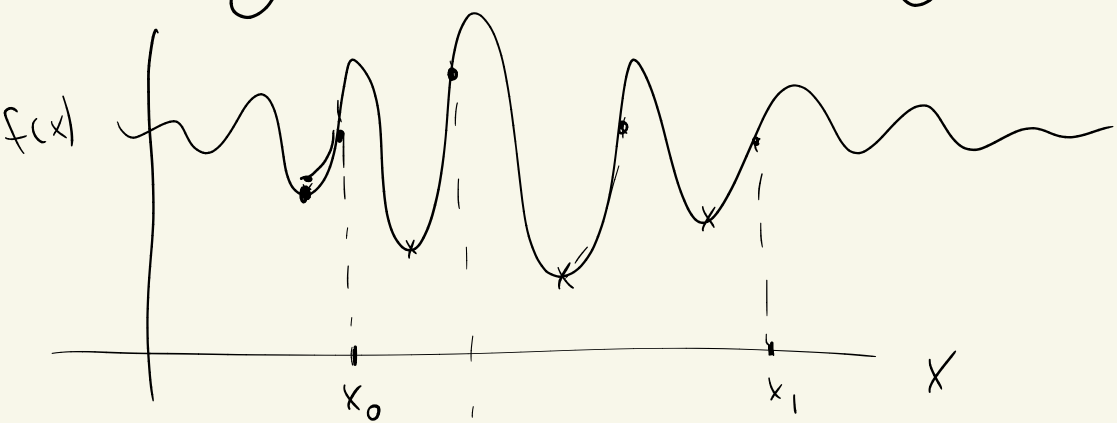
$$\delta\vec{x} = -\left[\frac{\partial^2 f}{\partial \vec{x}^2}\right]^{-1}_{\vec{x}_1} \cdot \left(\frac{\partial f}{\partial \vec{x}}\right)\Bigg|_{\vec{x}_1} \cdots\cdots \text{repeat}$$

- Require F to be invertible

- Steepest Descent ....

$\vec{x}_{n+1} = \vec{x}_n - \lambda \left.\frac{\partial f}{\partial x}\right|_{\vec{x}_n}$ → iterate ..... becomes slow as $\left.\frac{\partial f}{\partial x}\right|_{x_n} \to 0$ ..

Necessary Condition Methods only give us local solutions.



General Approach ....

- Sample a broad range of the domain
- Search and local optima
- Iterate till convergence .....

→ Genetic Algorithms

→ SNOPT, others

Parametric Optimization with Constraints

$f(\vec{x})$ continuous over a compact set R

Find extremal values of $J = f(\vec{x})$ subject to constraints

$$g_{\hat{j}}(\vec{x}) = 0 \quad ; \quad \hat{j} = 1, 2, \ldots, m \quad \Rightarrow \quad \vec{g}(\vec{x}) \in \mathbb{R}^m$$

IF $m > n \Rightarrow$ overconstrained problem, do not have any "choice" in J, solution may not even exist.

$m = n \Rightarrow$ For independent set of $g_{\hat{j}}(\vec{x})$, solution will exist in general, but potentially with limited choice in optimal values J.

Ex: $J = x$

$g(x) = x^2 - 1 \Rightarrow x = \pm 1$

possible values $J = +, -1$ .... discrete set to chose from.

$m < n \Rightarrow$ possible to satisfy constraints <u>and</u> have a choice to minimize J.

Reduction & Elimination

- Solve $g_j(x_1, x_2, \dots, x_n) = 0 \quad j = 1, 2, \dots, m$

these for states $x_{n-m+1}, x_{n-m+2}, \dots, x_n$ in terms of $(x_1, x_2, \dots, x_{n-m})$. Assume g_j are lin. indep.

Get $\boxed{y_i(x_1, x_2, \dots, x_{n-m}) = x_{n-m+i} \; ; \; i = 1, 2, \dots, m}$

Rewrite the problem as:

$$J = f(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_{n-m}, y_1, y_2, \dots, y_m)$$

Only free variables are $x_1, x_2, \dots, x_{n-m}$ as $y_i(x_1, x_2, \dots, x_{n-m})$.

Nec Conds:

$$\boxed{\frac{\partial J}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial x_i} = 0 \qquad i = 1, 2, \dots, n-m}$$

But... we may not be able to solve for the states....
" " " want to " " " "

$J = f(\vec{x})$, $\vec{g}(\vec{x}) = 0$ $\qquad$ $\vec{g} \in \mathbb{R}^m$ $\qquad$ $n > m$

$\vec{x} \in \mathbb{R}^n$

Given a trial point $\vec{x}_0 \mid \vec{g}(\vec{x}_0) = 0$,

want to find a neighboring point such that ....

$$J_0 + \delta J = f(\vec{x}_0 + \delta\vec{x}), \quad \boxed{\delta J = \left.\frac{\partial f}{\partial \vec{x}}\right|_0 \cdot \delta\vec{x} + \cdots}$$

and subject to

$$\vec{g}(\vec{x}_0 + \delta\vec{x}) = 0 = \vec{g}(\vec{x}_0) + \left.\frac{\partial \vec{g}}{\partial \vec{x}}\right|_0 \cdot \delta\vec{x} + \cdots$$

To be optimal $\delta J = 0 = \left.\frac{\partial f}{\partial \vec{x}}\right|_0 \cdot \delta\vec{x}$, we can't choose $\delta\vec{x}$ arbitrarily!

$\delta\vec{x}$ is subject to $0 = \left.\frac{\partial g}{\partial \vec{x}}\right|_0 \cdot \delta\vec{x}$ Need to satisfy <u>both</u> simultaneously

The $\delta\vec{x}$ are not lin. indep. Any variation $\delta J = \frac{\partial F}{\partial \vec{x}}\Big|_0 \cdot \delta\vec{x}$ that lies along $\left(\frac{\partial \vec{g}}{\partial \vec{x}}\right)\Big|_0 \cdot \delta\vec{x} = 0$ must be nulled out.

We can acheive this by adding "m" indep. parameters ...

$$\left[\frac{\partial F}{\partial \vec{x}}\Big|_0 \cdot \delta\vec{x} + \vec{\lambda} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}\Big|_0 \cdot \delta\vec{x} \right] = 0 \ ;$$

$$\underbrace{\left[\frac{\partial F}{\partial \vec{x}}\Big|_0 + \vec{\lambda} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}\Big|_0 \right]}_{0} \cdot \delta\vec{x} = 0 \Rightarrow \text{allows the } \delta\vec{x} \text{ to be } \underline{\text{free}}$$

0 n conditions, $\vec{x}, \vec{\lambda} \Rightarrow n+m$ indep. variables

$\vec{g}(\vec{x})$, m constraint conditions

Introduce Lagrange Multipliers $\vec{\lambda} \in \mathbb{R}^m$

Introduce the "Lagrangian" cost function

$$\boxed{L(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \vec{\lambda} \cdot \vec{g}(\vec{x})}$$

Lawden calls this the "Hamiltonian"

$$= f(\vec{x}) + \lambda_1 g_1(\vec{x}) + \lambda_2 g_2(\vec{x}) + \cdots + \lambda_m g_m(\vec{x})$$

Nec. Conds....

$$\frac{\partial L}{\partial \vec{x}} = \vec{0} \in \mathbb{R}^n \Rightarrow \frac{\partial F}{\partial \vec{x}} + \underset{\sim\sim\sim\sim}{\vec{\lambda} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}} = 0 \qquad n$$

$$\frac{\partial L}{\partial \vec{\lambda}} = \vec{0} \in \mathbb{R}^m \Rightarrow \vec{g}(\vec{x}) = \vec{0} \qquad + \ m$$

$\vec{x}, \vec{\lambda} \Rightarrow n+m$ variables

$n+m$ conditions

Must assume that $\frac{\partial g_j}{\partial x_i}$ are all lin. Indep.

Else $\exists \lambda_j \ni$

$$\sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0$$

How does this work?

$J = f(x,y)$; $x, y \in \mathbb{R}$

$g(x,y) = 0$

Rewritten nec. conds indep. of finding $y(x)$ function

Reduc. & Elimination

$y = y(x) \Rightarrow \boxed{J = f(x, y(x))}$

Staring values... $\dfrac{df}{dx} = \boxed{\dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial x} = 0}$

We also have $\boxed{g(x, y(x)) = 0} \Rightarrow \dfrac{dg}{dx} = \boxed{\dfrac{\partial g}{\partial x} + \dfrac{\partial g}{\partial y}\dfrac{\partial y}{\partial x} = 0}$

$$\frac{\partial y}{\partial x} = \frac{-\left(\frac{\partial g}{\partial x}\right)}{\left(\frac{\partial g}{\partial y}\right)} \Rightarrow \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\left[\frac{-\left(\frac{\partial g}{\partial x}\right)}{\left(\frac{\partial g}{\partial y}\right)}\right] = 0 \Rightarrow \boxed{\frac{\partial g}{\partial y}\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} = 0}$$

$\boxed{g(x,y) = 0}$

Lagrangian Approach

$$L = f(x,y) + \lambda g(x,y)$$

$$\boxed{\begin{aligned} \frac{\partial L}{\partial x} &= \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial L}{\partial y} &= \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} &= g(x,y) = 0 \end{aligned}}$$

Solve for λ

$$\boxed{\lambda = \frac{-\frac{\partial f}{\partial x}}{\left(\frac{\partial g}{\partial x}\right)} = \frac{-\frac{\partial f}{\partial y}}{\left(\frac{\partial g}{\partial y}\right)}}$$

$$\Downarrow$$

$$\boxed{\frac{\partial f}{\partial y}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} = 0}$$

$$\boxed{g(x,y) = 0}$$

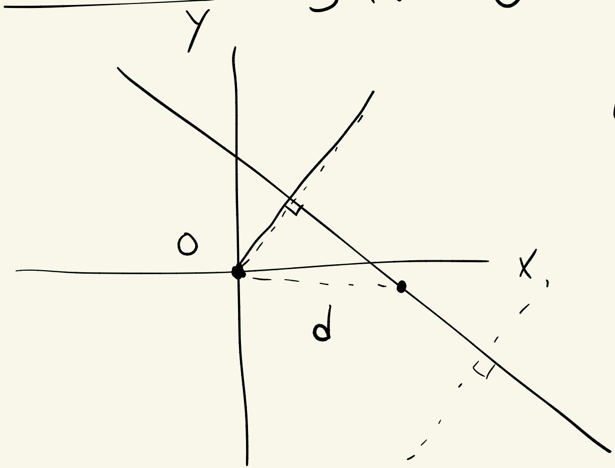
Same Conditions

$$\lambda \neq \frac{dy}{dx}$$

But.... $\boxed{\lambda \simeq -\frac{\partial f}{\partial g}}$ ☆ Change in cost over change in Constraint

See Lawden 1.4

Ex 2 Shortest distance from a point to a line.



Line : $\boxed{ax + by + c = 0}$

Solve for

$$\boxed{y = -\frac{a}{b}x - \frac{c}{b}}$$

Solution $\Rightarrow$ take ~~table~~ $\perp$ to the line,

$y_\perp = \frac{b}{a}x + d$; $\Rightarrow$ $\boxed{y_\perp = \frac{b}{a}x}$

Need $y_\perp = 0$ at $x = 0 \Rightarrow d = 0$

Optimal point is defined by

$$y = \boxed{-\frac{a}{b}x^* - \frac{c}{b} = \frac{b}{a}x^*} = y_\perp = y^*$$

$$\boxed{x^* = \frac{-ac}{a^2+b^2} \; ; \; y^* = \frac{-bc}{a^2+b^2}}$$

Use $J = d^2 = x^2 + y^2$

$g(x) = ax + by + c = 0$

$$\boxed{L = x^2 + y^2 + \lambda(ax + by + c)}$$

Nec. Cnds

$$\frac{\partial L}{\partial x} = 2x + a\lambda = 0$$

$$\frac{\partial L}{\partial y} = 2y + b\lambda = 0$$

$\Rightarrow$

$$\boxed{x = \frac{-a\lambda}{2}, \quad y = -\frac{b\lambda}{2}}$$

$$\frac{\partial L}{\partial \lambda} = ax + by + c = 0$$

$\longrightarrow$

$$\boxed{-\frac{a^2\lambda}{2} - \frac{b^2\lambda}{2} + c = 0} \Rightarrow$$

$$\boxed{\lambda = \frac{2c}{a^2+b^2}}$$

$$\boxed{x^* = \frac{-ac}{a^2+b^2}, \quad y^* = \frac{-bc}{a^2+b^2}}$$

$$J = \boxed{d = \sqrt{x^2+y^2}} \quad ; \quad g(x) = ax + by + c = 0$$

$$L = \sqrt{x^2+y^2} + \lambda(ax+by+c)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= \frac{x}{\sqrt{x^2+y^2}} + \lambda a = 0 \\ \frac{\partial L}{\partial y} &= \frac{y}{\sqrt{x^2+y^2}} + \lambda b = 0 \end{aligned}\right\} \qquad \left.\begin{aligned} x &= -\lambda a d \\ y &= -\lambda b d \end{aligned}\right\} \Rightarrow$$

$$\frac{\partial L}{\partial \lambda} = ax + by + c = 0$$

$$x^2 + y^2 = \cancel{d^2} = \lambda^2 a^2 d^2 + \lambda^2 b^2 d^2$$

$$= \lambda^2 \cancel{d^2}(a^2+b^2)$$

$$\boxed{\lambda^2 = \frac{1}{a^2+b^2}} \Rightarrow \boxed{\lambda = \frac{1}{\sqrt{a^2+b^2}}}$$

$$-\lambda a^2 d - \lambda b^2 d + c = 0$$

$$\frac{c}{a^2+b^2} = \lambda d \Rightarrow \boxed{d^* = \frac{c}{\sqrt{a^2+b^2}}}$$

$$\boxed{\begin{aligned} x^* &= \frac{-ac}{a^2+b^2} \\ y^* &= \frac{-bc}{a^2+b^2} \end{aligned}}$$

Inequality Constraints

Min/Max $J = f(\vec{x})$ $\quad \vec{x} \subset R \in \mathbb{R}^n$

subject to $\boxed{g_j(\vec{x}) = 0}$ $\quad j = 1, \dots, m$

$h_k(\vec{x}) \geq 0$ $\quad k = 1, \dots, l$

$\boxed{l + m \leq n}$

Mathematical Programming Problem (MPP)

Many traj. opt. + opt. control problems are reducible to this form.

If we consider a point $\vec{x}_0 \in R \dots \ni h_k(\vec{x}_0) > 0 \Rightarrow$

Constraint k is <u>inactive</u>. If $h_k(\vec{x}_0) \leq 0 \Rightarrow$ constraint is <u>active</u> and must be shifted to enforce $\boxed{h_k(\vec{x}_0 + \delta\vec{x}) = 0.}$

Full set of Necessary Conditions

Fritz-John Conditions : $\vec{x}^*$ is the optimum of the MPP iff

Define : $L(\vec{x}, \vec{\lambda}, \vec{\sigma})$ $\quad \vec{x} \subset R \in \mathbb{R}^n$

$\vec{\lambda} \in \mathbb{R}^m$

$\vec{\sigma} \in \mathbb{R}^l$

$$L = \lambda_0 f(\vec{x}) + \sum_{i=1}^{m} \lambda_i g_i(\vec{x}) + \sum_{j=1}^{l} \sigma_i h_i(\vec{x})$$

$$\boxed{L = \lambda_0 F(\vec{x}) + \vec{\lambda} \cdot \vec{g}(\vec{x}) + \vec{\sigma} \cdot \vec{h}(\vec{x})}$$

N.C. : There exists ($\exists$) $(\lambda_0, \vec{\lambda}, \vec{\sigma}) \neq (0, \vec{0}, \vec{0})$; $\lambda_0 \geq 0$, $\sigma_k \leq 0$ $\quad k = 1, \dots, l$

such that :

$$\left.\frac{\partial L}{\partial \vec{x}}\right|_{\vec{x}^*} = \vec{0} \quad ; \quad \vec{g}(\vec{x}^*) = \vec{0}, \quad \vec{h}(\vec{x}^*) \geq 0$$

$\lambda_0 = 0$? Only occurs when the $\vec{g}(\vec{x})$ + active $\vec{h}(\vec{x})$ are not linearly independent.

If ($\frac{\partial \vec{g}}{\partial \vec{x}}$ + $\frac{\partial \vec{h}}{\partial \vec{x}}$ are all linearly independent) then we can set $\lambda_0 = 1 \Rightarrow$ conditions become the Karush, Kuhn-Tucker Conditions (KKT). The constraint assumption is called "constraint qualification".

We get the KKT Nec. Conditions by setting $\lambda_0 = 1$ in the FJ conditions, assuming constraint qualification.

Why is $\sigma_k \leq 0$.

Define the set of all "active" inequality constraints for a given point $\vec{x}_0$.

$$I(\vec{x}) = \{ \hat{i} \mid h_i(\vec{x}) = 0 \}$$

If $j \notin I(\vec{x})$, then the constraint is <u>inactive</u> and

$\boxed{\sigma_{\hat{j}} = 0}$.

If $\hat{j} \in I(\vec{x})$, why is $\sigma_j < 0$ ____ ____

Assume we identify an optimal value $\vec{x}^*$ with no ineq. constraints active:

$$\left.\frac{\partial L}{\partial \vec{x}}\right|_{\vec{x}^*} = \left.\frac{\partial f}{\partial \vec{x}}\right|_{*} + \vec{\lambda} \cdot \left.\frac{\partial \vec{g}}{\partial \vec{x}}\right|_{*} = 0 \quad (\text{Assume } \vec{g}(\vec{x}^*) = \vec{0})$$

Consider the Lagrangian at this point.

$$L^* = f(\vec{x}^*) + \vec{\lambda} \cdot \cancel{\vec{g}(\vec{x}^*)} = F(\vec{x}^*)$$

If we add/active an ineq. constraint, force a shift in the optimal point from $\vec{x}^* \Rightarrow \vec{x}^* + \delta\vec{x}^* \Rightarrow h_1(\vec{x}^*) < 0$

$h_1(\vec{x}^* + \delta\vec{x}) = 0$

$$\boxed{L^*(\vec{x}^* + \delta\vec{x}^*) = f(\vec{x}^* + \delta\vec{x}^*) + \vec{\lambda} \cdot \vec{g}(\vec{x}^* + \delta\vec{x}^*) + \sigma_1 h_1(\vec{x}^* + \delta\vec{x}^*) \geq L(\vec{x}^*)}$$

Constraint Principle: Cost can only increase with the imposition of constraints.

$$L^*(\vec{x}^* + \delta\vec{x}^*) = \cancel{F(\vec{x}^*)} + \cancel{\left.\frac{\partial f}{\partial \vec{x}}\right|_{x^*} \cdot \delta\vec{x}}^{\,0} + \cancel{\vec{\lambda} \cdot g(\vec{x}^*)}^{\,0} + \cancel{\vec{\lambda} \cdot \left.\frac{\partial g}{\partial \vec{x}}\right| \cdot \delta\vec{x}}^{\,0}$$

$$+ \sigma_1 h(\vec{x}^*) + \cancel{\sigma_1 \left.\frac{\partial h}{\partial \vec{x}}\right|_* \cdot \delta\vec{x}} \geq \cancel{F(\vec{x}^*)} + \cancel{\vec{\lambda} \cdot \vec{g}(\vec{x}^*)}^{\,0}$$

$\sigma_1 h(\vec{x}^*) \geq 0$; $\Rightarrow \boxed{\sigma_1 \leq 0}$

Local & Global Sufficiency Conditions

Local: Start w/ the cost Lagrangian

$$L(\vec{x},\vec{\lambda},\vec{\sigma}) = f(\vec{x}) + \vec{\lambda}\cdot\vec{g}(\vec{x}) + \vec{\sigma}\cdot\vec{h}(\vec{x})$$

$$f \in \mathbb{R},\ \vec{x} \in \mathbb{R}^n,\ \vec{\lambda},\vec{g} \in \mathbb{R}^m,\ \vec{\sigma},\vec{h} \in \mathbb{R}^\ell$$

- Find the 2nd variation of L, subject to all constraints, assuming that the KKT N.C. are satisfied.

Need to look/consider all admissible variations.

$$\delta_{\begin{pmatrix}\vec{x}\\\vec{\lambda}\\\vec{\sigma}\end{pmatrix}} L \Big|_{*} = 0 \text{ , specifically}$$

$$\delta_{\vec{x}} f + \vec{\lambda}^* \cdot \delta_{\vec{x}} \vec{g} + \vec{\sigma}^* \cdot \delta_{\vec{x}} \vec{h} = 0$$

$$\left[\frac{\partial f}{\partial \vec{x}}\Big|_* + \vec{\lambda}^* \cdot \frac{\partial \vec{g}}{\partial \vec{x}}\Big|_* + \vec{\sigma}^* \cdot \frac{\partial \vec{h}}{\partial \vec{x}}\Big|_* \right] \cdot \delta \vec{x} = 0$$

$$\delta_{\vec{\lambda}} L = \delta \vec{\lambda} \cdot \vec{g}(\vec{x}) \Big|_* = 0$$

$$\delta_{\vec{\sigma}} L = \delta \vec{\sigma} \cdot \vec{h}(\vec{x}) \Big|_* = 0$$

Consider the 2nd Variation of L evaluated at $\vec{x}^*, \vec{\lambda}^*, \vec{\sigma}^*$

– Take the Variation first, then insert the Nec. Cond.

Notation: $\delta_x F = \frac{\partial F}{\partial \vec{x}} \cdot \delta \vec{x}$, $\delta_{\vec{x}}^2 F = \delta_x \left(\frac{\partial F}{\partial \vec{x}} \cdot \delta \vec{x} \right) = \delta \vec{x} \cdot \left[\frac{\partial^2 F}{\partial \vec{x}^2} \right] \cdot \delta \vec{x} = \delta \vec{x}^T \left[\frac{\partial^2 F}{\partial \vec{x}^2} \right] \delta \vec{x}$

Note $\delta^2 \vec{x} = \delta(\delta \vec{x}) = 0 \implies \frac{\partial^2 F}{\partial \vec{x}^2} \in \mathbb{R}^{n \times n}$

$$\delta_{\vec{x}}^2(\vec{\lambda}\cdot\vec{g}) = \delta_{\vec{x}}\left(\vec{\lambda}\cdot\delta\vec{g}\right) = \delta_{\vec{x}}\left(\vec{\lambda}\cdot\frac{\partial\vec{g}}{\partial\vec{x}}\cdot\delta\vec{x}\right) = \underbrace{\delta\vec{x}\cdot\left(\vec{\lambda}\cdot\frac{\partial^2\vec{g}}{\partial\vec{x}^2}\right)\cdot\delta\vec{x}}$$

$$\vec{g}\in\mathbb{R}^m$$

$$\frac{\partial\vec{g}}{\partial\vec{x}}\in\mathbb{R}^{m\times n} \qquad \frac{\partial^2\vec{g}}{\partial\vec{x}^2}\in\mathbb{R}^{m\times n\times n} \quad ; \quad \vec{\lambda}\cdot\frac{\partial^2\vec{g}}{\partial\vec{x}^2}\in\mathbb{R}^{n\times n}$$

We will seperately consider variations in $\delta\vec{x}, \delta\vec{\lambda}, \delta\vec{\sigma}$

$$\delta\vec{x}: \left[\frac{\partial f}{\partial\vec{x}} + \vec{\lambda}\cdot\frac{\partial\vec{g}}{\partial\vec{x}} + \vec{\sigma}\cdot\frac{\partial\vec{h}}{\partial\vec{x}}\right]\cdot\delta\vec{x} = 0$$

$$\delta\vec{\lambda}: \quad \delta\vec{\lambda}\cdot\vec{g} = 0$$

$$\delta\vec{\sigma}: \quad \delta\vec{\sigma}\cdot\vec{h} = 0$$

Second variations

$$\delta^2_{\vec{\lambda}} L = \delta^2_{\vec{\sigma}} L = 0$$

$$\delta^2_{\vec{\lambda}\vec{x}} L = \delta_{\vec{x}}\left(\delta\vec{\lambda}\cdot\vec{g}\right) = \underline{\delta\vec{\lambda}\cdot\frac{\partial\vec{g}}{\partial\vec{x}}\cdot\delta\vec{x}} \neq 0$$

But.... once evaluated along the N. C., for an allowable variation $\delta\vec{x}$, we <u>must</u> have $\frac{\partial\vec{g}}{\partial\vec{x}}\cdot\delta\vec{x} = 0$, since we enforce the constraint function.

Same $\delta^2_{\vec{\sigma}\vec{x}} L = 0$.

Good news $\Rightarrow$ Only need to consider variations in $\delta\vec{x}$.

$$\delta^2 L = \delta^2_{\vec{x}\vec{x}} L = \delta\vec{x}\cdot\left[\frac{\partial^2 f}{\partial\vec{x}^2} + \vec{\lambda}\cdot\frac{\partial^2\vec{g}}{\partial\vec{x}^2} + \vec{\sigma}\cdot\frac{\partial^2\vec{h}}{\partial\vec{x}^2}\right]\Bigg|_{(\vec{x}^*,\lambda^*,\sigma^*)}\cdot\delta\vec{x} \geq 0$$

to be a local min value.

$$L_{\vec{x}\vec{x}}\Big|_{\vec{x}^*,\vec{\lambda}^*,\vec{\sigma}^*} = \left[\frac{\partial^2 f}{\partial \vec{x}^2} + \vec{\lambda}\cdot\frac{\partial^2 \vec{g}}{\partial \vec{x}^2} + \vec{\sigma}\cdot\frac{\partial^2 \vec{h}}{\partial \vec{x}^2} \right]\Bigg|_{(\vec{x}^*,\vec{\lambda}^*,\vec{\sigma}^*)}$$

pos. def or pos. semi-def. <u>Local Condition</u>

Global Min/Max Conditions can be found under <u>convexity</u> assumptions.

In the MPP assume the KKT conditions apply (constraint qualification).

Then if $f, \vec{g}, \vec{h}$ are C^1 (First derivative continuous)

$f, \vec{h}$ are convex functions

$\vec{g}$ is Affine ($= \bar{\bar{A}}\cdot\vec{x} + \vec{b} = \vec{0}$)

Then the constraint set is convex. And if $\vec{x}^*$ satisfies KKT

$\boxed{\text{! } \vec{x}^* \text{ is the global minimizer}}$

Multi-dimensional optimal cost functions.

What if we have p different cost functions that we want to optimize.

$$\vec{F}(\vec{x}) \in \mathbb{R}^p \qquad F_1, F_2, F_3, \dots, F_p$$

Two general approaches

Weighted cost function.

Choose relative weights between each cost function.

$$w_i\,;\; i=1,2,\dots,p \qquad \sum_{i=1}^{p} w_i = 1$$

Define the combined cost function

$$\boxed{F(\vec{x}, \vec{w}) = \sum_{i=1}^{p} w_i f_i(\vec{x})}$$

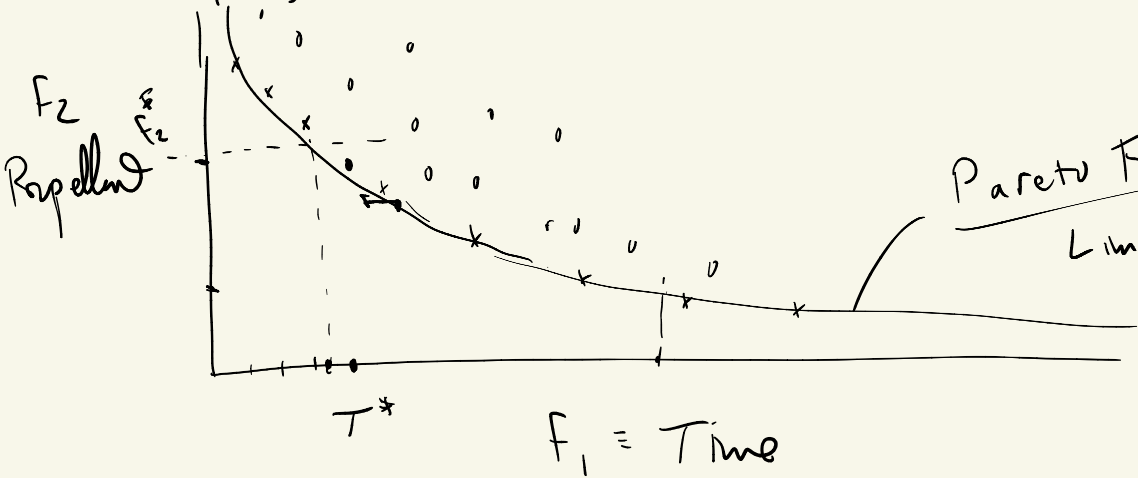
For a fixed set of weights $\Rightarrow$ we get back to the usual case.

Optimal value + solution are then

$$\boxed{\vec{x}^*(\vec{w}), \quad F^*(x^*(\vec{w}), \vec{w})}$$

Another approach .... Consider all the cost functions at once and balance the different costs. Consider a 2-D case

$F_1(\vec{x}),\ F_2(\vec{x}) \Rightarrow$ Time, Propellant



Pareto Front

Limit between the costs such that the reduction in one cost always leads to an increase in the other cost.

$r < 11.6$