We have so far seen:

$$\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & x \in \mathbb{R}^n \end{cases} \qquad \begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & h_1(x) = 0 \\ & \vdots \\ & h_m(x) = 0 \end{cases}$$

Now we will consider inequality constraints:

$$\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{cases}$$

How do we solve this type of a problem?

K-K-T conditions. $\underbrace{\text{Karush}}_{\text{Master's thesis 1939}} - \underbrace{\text{Kuhn - Tucker}}_{1951}$

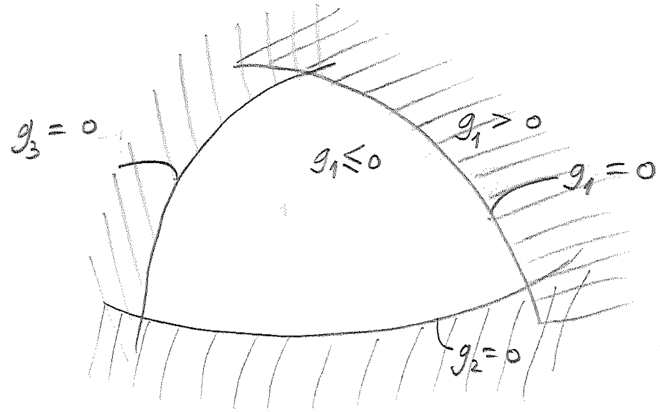
Minima of Functions of Several Variables with Inequalities as Side Constraints.
Univ. of Chicago

Nonlinear programming
Proceedings of the 2nd Berkeley Symposium

A mechanical analogue

Imagine a potential in $\mathbb{R}^2$ given by f.

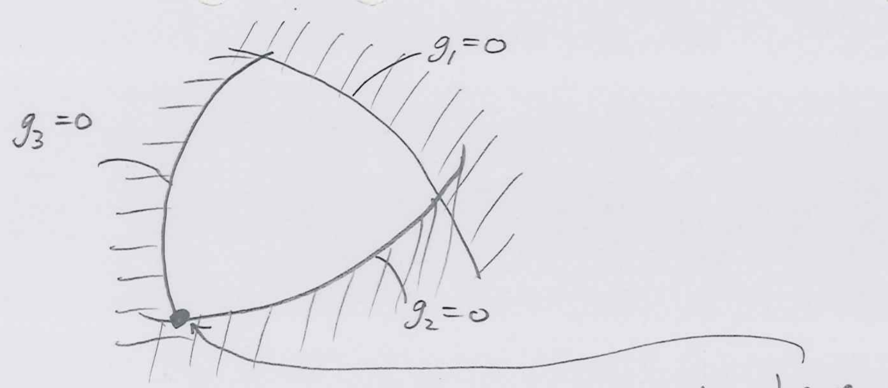
Suppose a particle is confined by walls described by
$$\begin{aligned} g_1(x) &> 0 \\ &\vdots \\ g_m(x) &> 0. \end{aligned}$$



Particle tries to minimize the potential energy.

So the particle tries to solve

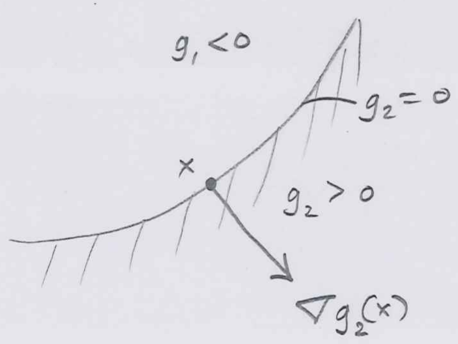
$$\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0. \end{cases}$$



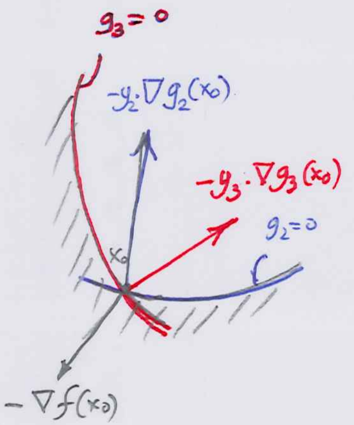
Suppose the particle comes to rest here.

The force on the particle is $-\nabla f(x_0)$

and this must be balanced by the normal forces from the walls.

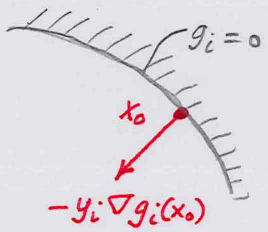
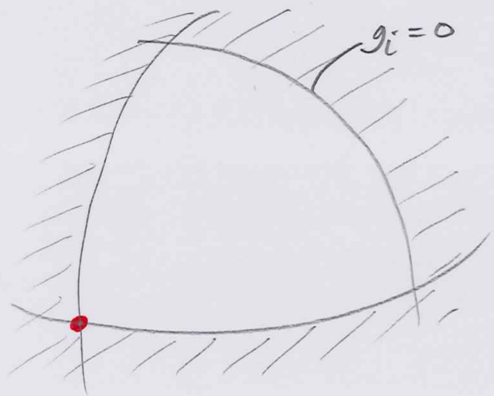


So normal force on the particle is $-y \cdot \nabla g_2(x)$ with $y>0$.



"Active" walls are wall 2 and wall 3, so we get

$$-\nabla f(x_0) - y_2 \nabla g_2(x_0) - y_3 \nabla g_3(x_0) = 0 . \quad \text{(Balance of forces)}$$

Two possible cases:		Normal force on particle from wall	$y_i g_i(x_0)$
Particle is touching the ith wall		$-y_i \nabla g_i(x_0)$ for some $y_i \geq 0$	$y_i \cdot \underbrace{g_i(x_0)}_{=}$ $= y_i \cdot 0$ $= 0$
Particle is not touching the ith wall		$0 = -y_i \nabla g_i(x_0)$ with $y_i := 0$	$\underbrace{y_i}_{=} \cdot \underbrace{g_i(x_0)}_{<0}$ $= 0 \cdot \boxed{*}$ $= 0.$

So we arrive at the following KKT - conditions:

The optimal x_0 is s.t. $\exists\, y_1, \dots, y_m$ s.t

(KKT-1) $\nabla f(x_0) + y_1 \nabla g_1(x_0) + \dots + y_m \nabla g_m(x_0) = 0$ (Force balance.)

(KKT-2) $\left.\begin{array}{l} g_1(x_0) \leq 0 \\ \vdots \\ g_m(x_0) \leq 0 \end{array}\right\}$ (Particle is within the region enclosed by the walls.)

(KKT-3) $\left.\begin{array}{l} y_1 \geq 0 \\ \vdots \\ y_m \geq 0 \end{array}\right\}$ (The walls exert normal forces.)

(KKT-4) $\left.\begin{array}{l} y_1\, g_1(x_0) = 0 \\ \vdots \\ y_m\, g_m(x_0) = 0 . \end{array}\right\}$ (Only those walls which the particle touches can exert nontrivial normal forces.)

Theorem. (Necessity of K-K-T conditions)

Consider the problem (P):
$$\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0. \end{cases}$$

If x_0 is a local minimizer for (P) and x_0 is a "regular" point,

then $\exists\, y \in \mathbb{R}^m$ s.t.

(KKT-1) $\nabla f(x_0) + y_1 \nabla g_1(x_0) + \dots + y_m \nabla g_m(x_0) = 0$

(KKT-2) $g_i(x_0) \leq 0, \quad i = 1, \dots, m.$

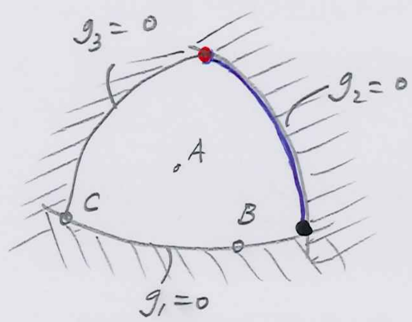
(KKT-3) $y_i \geq 0, \quad i = 1, \dots, m.$

(KKT-4) $y_i g_i(x_0) = 0, \quad i = 1, \dots, m.$

What does regular mean here?

"Active" index set of a feasible point: Given $x \in \mathcal{F}$, this is the set of is s.t. $g_i(x) = 0$.

Example:



For point A, $I_a(A) = \phi$

For point B, $I_a(B) = \{1\}$

For point C, $I_a(C) = \{1, 3\}$.

This is useful to keep track of where the point is in the feasible set.

Example: If $I_a(x) = \{2,3\}$, we know the point x is the red dot shown in the figure above.

If $I_a(x) = \{2\}$, we know the point x is on the blue wall shown above, but is not the red dot or the black dot.

How to find out if a point $x_0 \in \mathfrak{Z}_e$ is regular?

Look at the vectors $\nabla g_i(x_0)$ for $i \in I_a(x_0)$.

There should be no linear combination of these vectors with non negative scalars giving 0 except for the trivial one. (with all scalars zero).

Thus there do not exist scalars v_i $(i \in I_a(x_0))$, not all zeros,

s.t. $\sum\limits_{i \in I_a(x_0)} v_i \nabla g_i(x_0) = 0$

and $v_i \geq 0 \quad \forall\, i \in I_a(x_0)$.

Note that if $\nabla g_i(x_0)$ $(i \in I_a(x_0))$ are linearly independent, then x_0 is regular.

Example

$$\begin{cases} \text{minimize} & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \geq 1 \\ & x_2 \geq 0 \end{cases}$$

$$\left.\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \end{cases}\right\} \quad \text{where} \quad \begin{aligned} f(x) &:= x_1^2 + x_2^2 \\ g_1(x) &:= 1 - x_1 - x_2 \\ g_2(x) &:= -x_2 . \end{aligned}$$

$$\nabla f(x) = [2x_1 \quad 2x_2]$$

$$\left.\begin{aligned} \nabla g_1(x) &= [-1 \quad -1] \\ \nabla g_2(x) &= [0 \quad -1] \end{aligned}\right\} \text{ independent } \Rightarrow \text{ every feasible point is regular.}$$

(KKT-1)

$$2x_1 - y_1 = 0$$
$$2x_2 - y_1 - y_2 = 0$$

(KKT-2)

$$x_1 + x_2 \geq 1$$
$$x_2 \geq 0$$

(KKT-3)

$$y_1 \geq 0$$
$$y_2 \geq 0$$

(KKT-4)

$$y_1 \cdot (1 - x_1 - x_2) = 0$$
$$y_2 \cdot (-x_2) = 0 .$$

$\underline{1°}$ $\boxed{\begin{matrix} y_1 > 0 \\ y_2 > 0 \end{matrix}}$

<u>KKT-4</u>

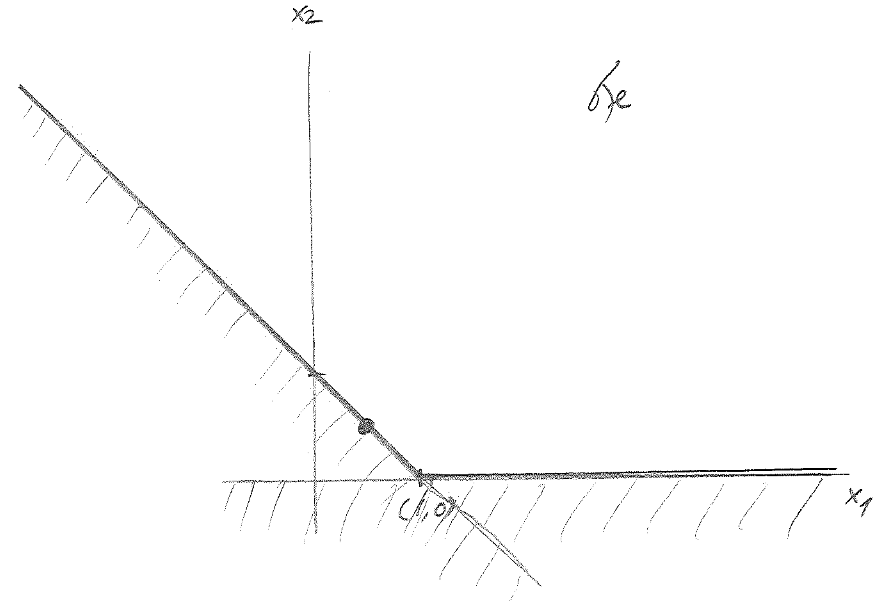
$$\left.\begin{matrix} x_1 + x_2 = 1 \\ x_2 = 0 \end{matrix}\right\} \Rightarrow \begin{matrix} x_1 = 1 \\ x_2 = 0 \end{matrix} \qquad x = (1, 0)$$

<u>KKT-1</u>

$$y_1 = 2x_1 = 2 > 0$$

$$y_2 = 2x_2 - y_1 = 0 - 2 = -2 < 0,$$ violates KKT-3.

So no KKT points (x, y) in this case.



$\underline{2°}$ $\boxed{\begin{matrix} y_1 > 0 \\ y_2 = 0 \end{matrix}}$

<u>KKT-4</u>

$x_1 + x_2 = 1$

<u>KKT-1</u>

$$\left.\begin{matrix} 2x_1 - y_1 = 0 \\ 2x_2 - y_1 = 0 \end{matrix}\right\} \Rightarrow x_1 = x_2$$

$$\Rightarrow x_1 = x_2 = \frac{1}{2}$$

$$y_1 = 2x_1 = 1 > 0.$$

$\exists$ a KKT point given by $x = (\frac{1}{2}, \frac{1}{2})$, $y = (1, 0)$.

$\underline{3°}$ $\boxed{\begin{matrix} y_1 = 0 \\ y_2 > 0 \end{matrix}}$

<u>KKT-4</u>

$x_2 = 0$

<u>KKT-1</u>

$2x_1 = 0$

$2x_2 - y_2 = 0$

$\Rightarrow x_1 = x_2 = 0$. But then $x_1 + x_2 = 0 \neq 1$; so KKT-2 is violated

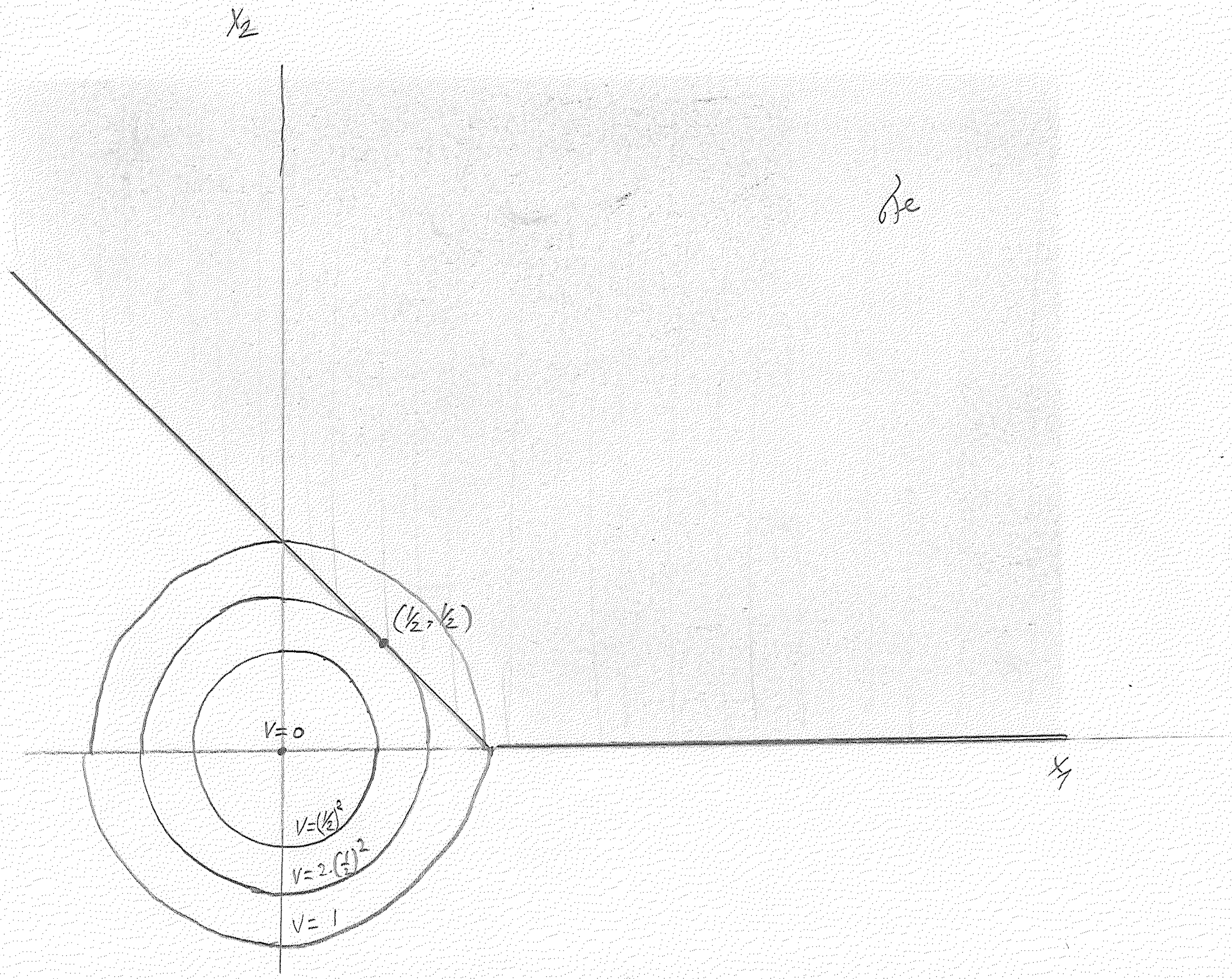
So no KKT points in this case.

$\underline{4°}$ $\boxed{\begin{matrix} y_1 = 0 \\ y_2 = 0 \end{matrix}}$

<u>KKT-1</u>

$$\left.\begin{matrix} 2x_1 = 0 \\ 2x_2 = 0 \end{matrix}\right\rangle \Rightarrow x_1 = x_2 = 0.$$ Again KKT-2 is violated

So no KKT points in this case.



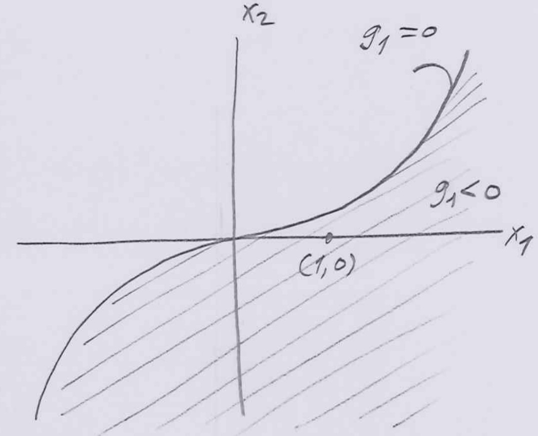
x_2
$\mathcal{F}_e$
$(1/2, 1/2)$
$V=0$
$V=(1/2)^2$
$V=2\cdot(\frac{1}{2})^2$
$V=1$
x_1

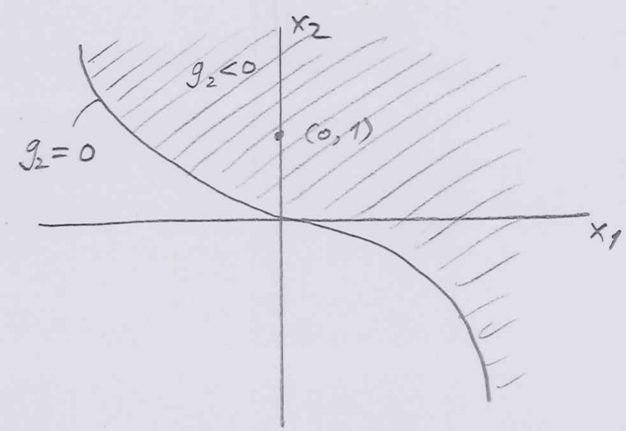
<u>Example</u> (Importance of the regularity condition)

Consider the problem
$$\begin{cases} \text{minimize} & x_1 \\ \text{s.t.} & x_2 \leq x_1^3 \\ & x_2 \geq -x_1^3 . \end{cases}$$

i.e.
$$\begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \end{cases}$$

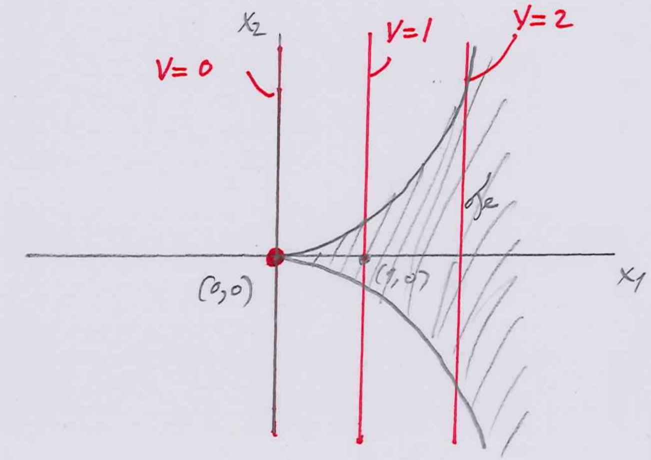
where
$$f(x) := x_1$$
$$g_1(x) := x_2 - x_1^3$$
$$g_2(x) := -x_1^3 - x_2 .$$





Clearly $(0,0)$ is a global minimizer.

<u>But</u> the KKT-conditions are <u>not</u> satisfied!



$\nabla f(x) = [1 \quad 0]$ So $\nabla f(0) = [1 \quad 0]$

$\nabla g_1(x) = [-3x_1^2 \quad 1]$ $\nabla g_1(x) = [0 \quad 1]$

$\nabla g_2(x) = [-3x_1^2 \quad -1]$. $\nabla g_2(x) = [0 \quad -1]$.

There is clearly no y_1, y_2 s.t. $\nabla f(0) + y_1 \nabla g_1(0) + y_2 \nabla g_2(0) = 0$

(since $\underbrace{\nabla f(0) + y_1 \nabla g_1(0) + y_2 \nabla g_2(0)}_{=[1 \quad y_1 - y_2]} \neq [0 \quad 0]$).

So KKT-1 can never be satisfied!

<u>What went wrong?</u>: The point $(0,0)$ is not regular.

Indeed, $1 \cdot \nabla g_1(0,0) + 1 \cdot \nabla g_2(0,0)$

$= 1 \cdot [0 \quad 1] + 1 \cdot [0 \quad -1]$

$= [0 \quad 0]$.

Sketch of the proof (of the theorem on necessity of the KKT-conditions).

Lemma. If x_0 is a local minimizer,

then there is no $d \in \mathbb{R}^n$ s.t. $\nabla f(x_0)\, d < 0$ and

$$\nabla g_i(x_0)\, d < 0 \quad \forall\ i \in I_a(x_0).$$

($\nabla g_i(x_0)\, d < 0 \ \ \forall\ i \in I_a(x_0)$ means we can start from x_0 and go inside $\mathcal{F}$ in the direction of d.

$\nabla f(x_0)\, d < 0$ means we can also reduce the cost.

Clearly this is impossible since x_0 is a local minimizer.)

Proof Suppose $\exists\ d$ s.t.

$$\nabla g_i(x_0)\, d < 0 \quad \forall\ i \in I_a(x_0)$$

Consider $\psi(t) = g_i(x_0 + td)$, $t \in \mathbb{R}$

$\psi'(t) = \nabla g_i(x_0 + td) \cdot d$.

$\psi'(0) = \nabla g_i(x_0) \cdot d < 0$.

So $\psi'(t) < 0$ for small ts.

We have

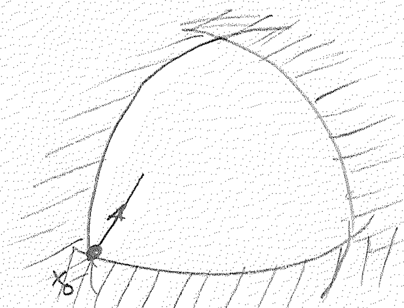
$$\begin{array}{ccccc} \psi(t) & - & \psi(0) & = & \underbrace{\int_0^t \psi'(\tau)\, d\tau}_{<0 \text{ for } t>0 \text{ small}} \\ \| & & \| & & \\ g_i(x_0+td) & - & g(x_0) & < & 0 \quad \text{for small } t>0 \end{array}$$

$$g_i(x_0+td) < g(x_0) \leq 0 \text{ for small } t>0 .$$

This is for is in $I_a(x_0)$

For other is, $g_i(x_0+td) < 0$ for small $t>0$ $(\because g_i(x_0) < 0)$.

So for <u>all</u> i, $g_i(x_0+td) < 0$ for small $t>0$



We know that

$$f(x) \geqslant f(x_0) \quad \text{for all } x \in \Omega \text{ near } x_0$$

In particular for all small $t>0$,

$$f(x_0+td) \geqslant f(x_0) .$$

So $\dfrac{f(x_0+td) - f(x_0)}{t} \geqslant 0$ for all small $t>0$.

$\nabla f(x_0) . d \geqslant 0$. This completes the proof. □

Farkas' lemma

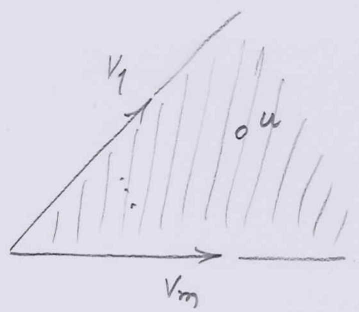
Let $v_1, \dots, v_m \in \mathbb{R}^n$,

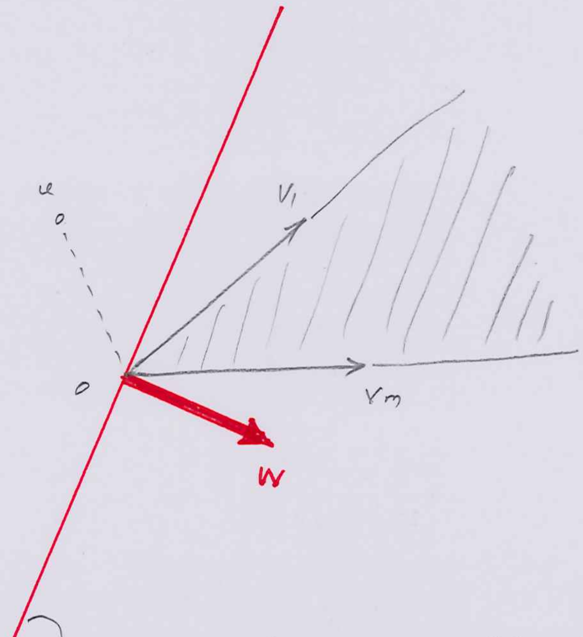
$u \in \mathbb{R}^n$.

Then one and only one of the following occurs:

(F1). $\exists\, w \in \mathbb{R}^n$ s.t. $w^T u < 0$ and $w^T v_1 \geq 0$, $\vdots$, $w^T v_m \geq 0$.

(F2) $\exists\ y_1 \geq 0, \dots, y_m \geq 0$ s.t. $u = y_1 v_1 + \dots + y_m v_m$





separating hyperplane $\{\xi \in \mathbb{R}^n : w^T \xi = 0\}$

$$\left.\begin{array}{l}\text{Lemma}\\ \quad + \\ \text{Farkas' lemma}\end{array}\right\}$$ gives the theorem on necessity of the KKT-conditions.

<u>Use of KKT-necessary conditions:</u>

Narrows down the possibilities for a local minimizer.

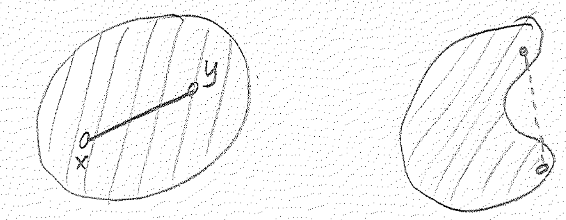
If we know that the problem has a global minimizer (for example using Weierstrass' theorem), then the KKT-necessary conditions are very useful.

Useful for ruling out points that are not local minimizers.

Convexity

convex set — convex function — convex problem ?

Convex set: $C \subset \mathbb{R}^n$ is convex if $\forall x, y \in C \quad \forall t \in (0,1)$,
$(1-t)x + ty \in C$.



Convex function: $f : C \longrightarrow \mathbb{R}$ is convex if
$\forall x, y \in C \quad \forall t \in (0,1) \qquad f((1-t)x + ty) \leq (1-t)f(x) + t f(y)$.

Theorem: If C has at least one interior point,
then $\boxed{f : C \longrightarrow \mathbb{R} \text{ is convex}} \Longleftrightarrow \boxed{\forall x \in C,\ F(x) \text{ is p.s.d.}}$

What is a convex problem?

Consider $(P): \begin{cases} \text{minimize} & f(x) \\ \text{s.t.} & x \in \mathcal{X} \end{cases}$

The problem (P) is called a convex problem if

(1) $\mathcal{X}$ is a convex set and

(2) $f: \mathcal{X} \to \mathbb{R}$ is a convex function.

Example.

(I) $(LP): \left. \begin{cases} \text{minimize} & c^T x \\ \text{s.t.} & Ax \leq b \end{cases} \right\}$ is a convex problem

Why? (1) $\mathcal{X} := \{ x \in \mathbb{R}^n : Ax \leq b \}$ is a convex set.

Let $x, y \in \mathcal{X}$. Then $\begin{matrix} Ax \leq b \\ Ay \leq b. \end{matrix}$

Let $t \in (0,1)$. Then $A((1-t)x + ty) = (1-t)Ax + tAy \leq (1-t)b + tb$
$= b.$

So $(1-t)x + ty \in \mathcal{X}$. Hence $\mathcal{X}$ is convex.

(2) $f(x) := c^T x$

$f: \mathcal{X} \longrightarrow \mathbb{R}$ is a convex function. For $x, y \in \mathcal{X}$ and $t \in (0,1)$, we have:

$f((1-t)x + ty) = c^T((1-t)x + ty) = (1-t)c^T x + t c^T y = (1-t)f(x) + t f(y)$!

So all linear programming problems are convex

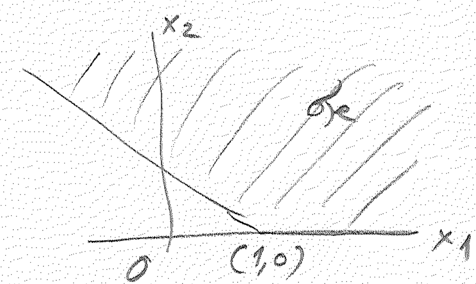
(II). $\begin{cases} \text{minimize} & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \geqslant 1 \\ & x_2 \geqslant 0 \end{cases}$ of the form $Ax \geqslant b$.

So $\mathcal{X}$ is convex.

Consider f given by $f(x) = x_1^2 + x_2^2$.

$\nabla f(x) = [2x_1 \quad 2x_2]$

$F(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is p.d. $\forall x \in \mathcal{X}$.

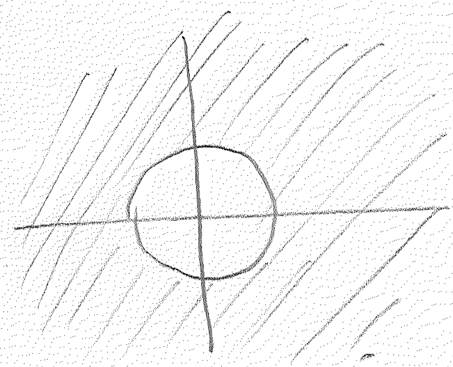


$\mathcal{X}$ has interior points

So f is convex.

(III) $\begin{cases} \text{minimize} & x_1 + x_2 \\ \text{s.t.} & x_1^2 + x_2^2 \geq 1 . \end{cases}$

is not a convex problem, since the feasible set is not convex



Take $x = (2, 0)$

$y = (-2, 0)$.

$t = \frac{1}{2}$

$(1-t)x + ty = \frac{1}{2}x + \frac{1}{2}y = \frac{1}{2}(0, 0)$

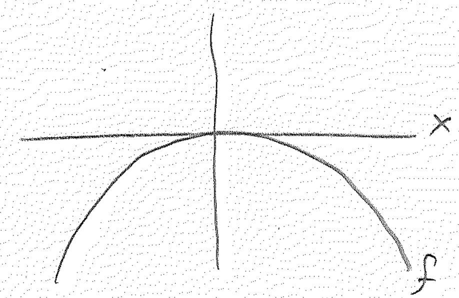
$= (0, 0)$

$\notin \mathcal{X}$.

(III) $\begin{cases} \text{minimize} & -x^2 \\ \text{s.t.} & x \in \mathbb{R} \end{cases}$

is not a convex problem, since although $\mathbb{R}$ is convex,

the function $x \mapsto -x^2$ is not a convex function



For example, take $x = -1$, $y = 1$, $t = 1/2$.

Then $f((1-t)x + ty) = f\left(-\frac{1}{2} + \frac{1}{2}\right) = f(0) = 0$.

While $(1-t)f(x) + t f(y) = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (-1) = -1$ and $0 \not\leq -1$.