Chapter 12: Theory of Constrained Optimization

Want:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases}$$

* f is the objective function
* $c_i, i \in \mathcal{E}$ are the equality constraints
* $c_i, i \in \mathcal{I}$ are the inequality constraints

Feasible set Ω:

$$\Omega = \{ x \mid c_i(x) = 0, \ i \in \mathcal{E}; \ c_i(x) \geq 0, \ i \in \mathcal{I} \}$$

or write:

$$\min_{x \in \Omega} f(x)$$

Local Solution: x^* is a local solution if $x^* \in \Omega$ and there is a neighborhood N of x^* such that: $f(x) \geq f(x^*)$ for $x \in N \cap \Omega$.

Strict local solution: x^* is a strict local solution if it is a local solution with $f(x) > f(x^*)$ for $x \in N \cap \Omega$, $x \neq x^*$, for some N of $x^* \in \Omega$

Isolated local minimizer: x^* is an isolated local minimizer if it is the **only one local minimizer** for $N(x^*)$,

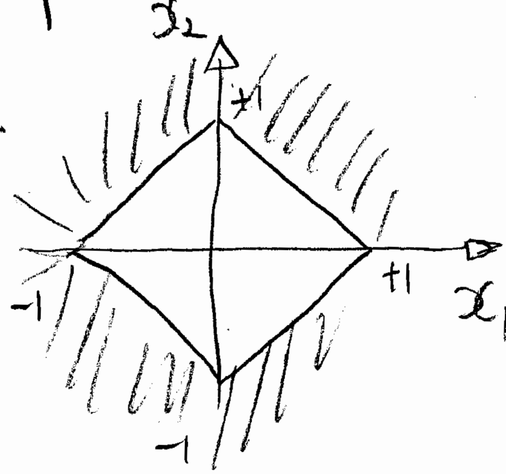
$x^* \in \Omega$, for all $x \in N \cap \Omega$.

Examples:

L_1-norm: $\|x\|_1 = |x_1| + |x_2| \leq 1$

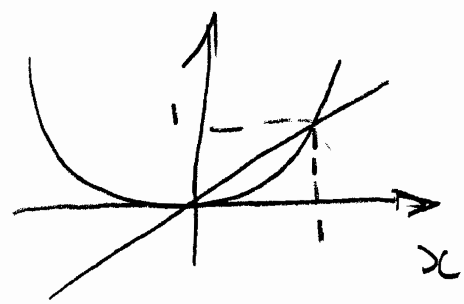
can also be described by:

$$\begin{aligned} x_1 + x_2 &\leq 1 \\ x_1 - x_2 &\leq 1 \\ -x_1 + x_2 &\leq 1 \\ -x_1 - x_2 &\leq 1 \end{aligned}$$



A second one:

$$f(x) = \max(x^2, x)$$



is the same as:

$$\min t \quad \text{s.t.} \begin{cases} t \geq x \\ t \geq x^2 \end{cases}$$

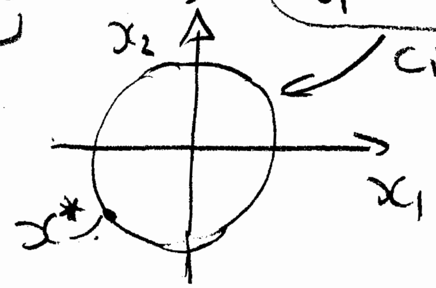
Active: if $c_i(x) = 0$, $i \in I$, then the inequality $c_i(x) \geq 0$ is said to be active. Else, it is inactive.

Ex 12.1

Consider: $\min x_1 + x_2$ s.t. $x_1^2 + x_2^2 - 2 = 0$

$$\underbrace{x_1^2 + x_2^2 - 2}_{c_1(x)} = 0 \Rightarrow x_1^2 + x_2^2 = (\sqrt{2})^2$$

circle.



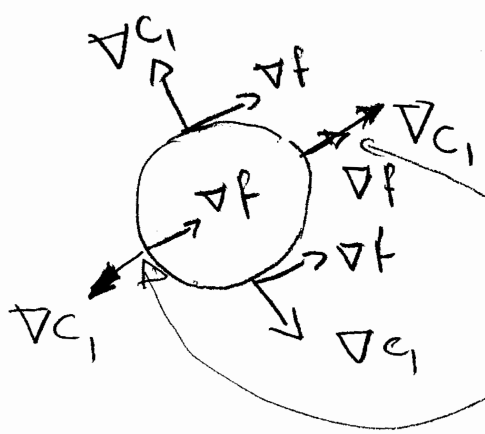
Solution is $(x_1, x_2) = (-1, -1)^T$ satisfying:

$$\boxed{\nabla f(x^*) = \lambda_1^* \nabla c_1(x^*)} \quad (*)$$

or:

$$\nabla f(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_1^* \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

is satisfied at x^* ($\lambda_1^* = -\frac{1}{2}$).



Note that there are two points where $(*)$ is satisfied (one for min, one for max)

To understand the conditions, note that $c_1(x) = 0$ within the feasible region.

$$\Rightarrow 0 = c_1(x+d) \approx c_1(x) + \nabla c_1(x)^T d$$
$$= 0 + \nabla c_1(x)^T d \approx 0$$

(*)

$$\Rightarrow \boxed{\nabla c_1(x)^T d = 0 \quad \text{(to first order app.)}}$$

Furthermore, for decreasing f:

$$0 > f(x+d) - f(x) \approx \nabla f(x)^T d$$

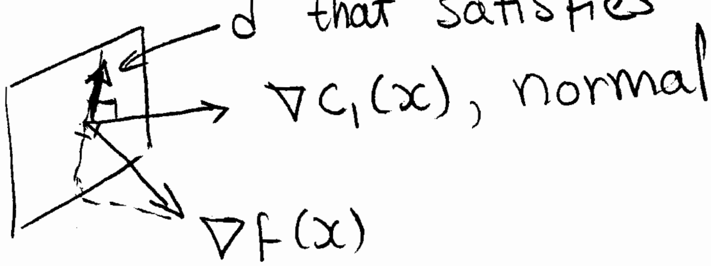
$$\Rightarrow \boxed{\nabla f(x)^T d < 0 \quad \text{(to first order app.)}}$$

(**)

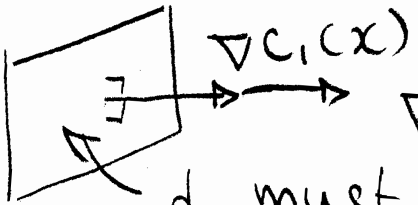
If there is a d that satisfies both, then by taking $x+d$, we can minimize f and stay feasible (up to first order):

$\Rightarrow$ for x^* optimal, we cannot find a d that satisfies both (*), (**). (necessary).

Now, if $\nabla c_1(x)$, $\nabla f(x)$ are not parallel, then, we can construct d by:



On the other hand, if $\nabla c_1(x), \nabla f(x)$ are parallel, no d can be found:



d must lie on this plane for ⊛, but if it does, (**) is violated: $\nabla f(x)^T d = 0$.

$$\Rightarrow \quad \boxed{\nabla f(x^*) = \lambda_1 \nabla c_1(x^*) \text{ at } x^* \quad \underline{\text{necessary}} \text{ condition.}}$$

Form the Lagrangian:

$$\mathcal{L}(x, \lambda_1) = f(x) - \lambda_1 c_1(x)$$

$$\Rightarrow \nabla_x \mathcal{L}(x, \lambda_1) = \nabla f(x) - \lambda_1 \nabla c_1(x)$$

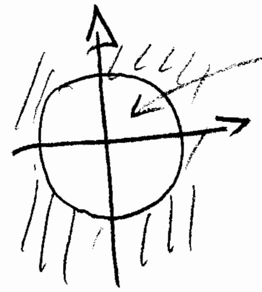
<u>Necessary condition</u>:

$$\exists \lambda_1^* \text{ such that: } \nabla_x \mathcal{L}(x^*, \lambda_1^*) = 0$$

However, it is not <u>sufficient</u>, and we cannot make it so by restricting the sign of λ_1.

A single inequality:

min $x_1 + x_2$ such that $2 - x_1^2 - x_2^2 \geq 0$.



inside of the circle.

Still, minimum at $(-1,-1)$

Again, consider:

$$0 \leq c_1(x+d) \approx c_1(x) + \nabla c_1(x)^T d$$

which requires:

$$\boxed{c_1(x) + \nabla c_1(x)^T d \geq 0}$$

*** up to first order app.

Also recall: $\boxed{\nabla f(x)^T d < 0}$, up to first order app.

① Consider an interior point, where

$$c_1(x) + \nabla c_1(x)^T d > 0,$$

for small-enough d, we have *** satisfied.

Set $d = -\nabla f(x)\, h$, h small.

Then this also satisfies: $\nabla f(x)^T d < 0$.

$\Rightarrow$ for <u>interior points</u>, $\nabla f(x) = 0$.

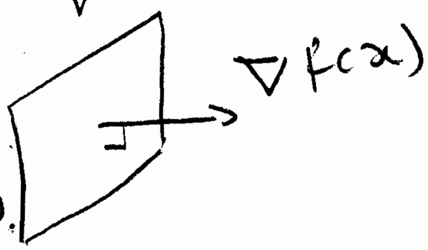
is required for x^*.

(II) On the boundary, we have:

$$\underbrace{\nabla f(x)^T d < 0}, \quad \underbrace{\nabla c_1(x)^T d \geq 0},$$

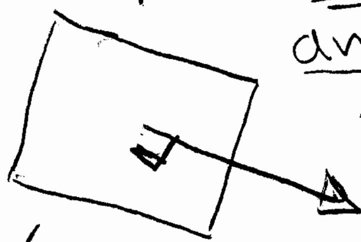
Behind plane:

$\nabla f(x)^T d < 0$



in-front and on-plane:

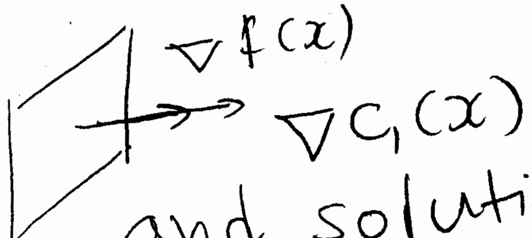
$\nabla c_1(x)^T d \geq 0$



Can always be combined to give a solution, unless $\boxed{\nabla f(x) = \lambda_1 \nabla c_1(x), \text{ for } \lambda_1 \geq 0.}$

If $\lambda_1 = 0 \Rightarrow \nabla f(x) = 0$, and we have a stationary point.

Else



and solutions have no points for d.

Lagrangian formulation

$$\boxed{\nabla_x \mathcal{L}(x^*, \lambda_1^*) = 0}, \quad \text{some } \lambda_1 \geq 0,$$

for $\mathcal{L}(x, \lambda_1) = f(x) - \lambda_1 c_1(x)$.

and $\boxed{\lambda_1^* c_1(x^*) = 0} \sim \Delta$

complimentarity condition.

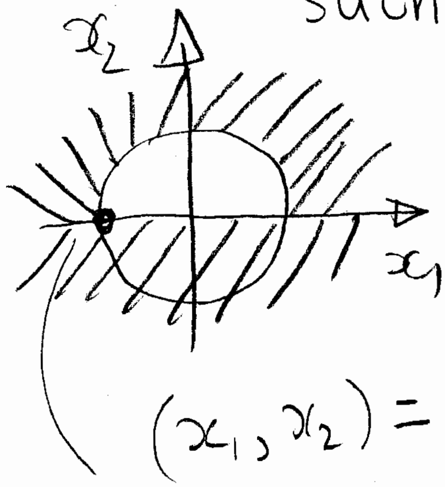
From Δ:

* minimum is at the boundary ($c_1(x^*) = 0$) OR
* it is a regular minimum with $\nabla f(x^*) = 0$ on the boundary.

Note that $\nabla f(x^*) = 0$ will satisfy Δ with $\lambda_1^* = 0$.

Two-inequality constraints

$$\min \; x_1 + x_2$$
such that: $\begin{cases} 2 - x_1^2 - x_2^2 \geq 0, \\ x_2 \geq 0 \end{cases}$



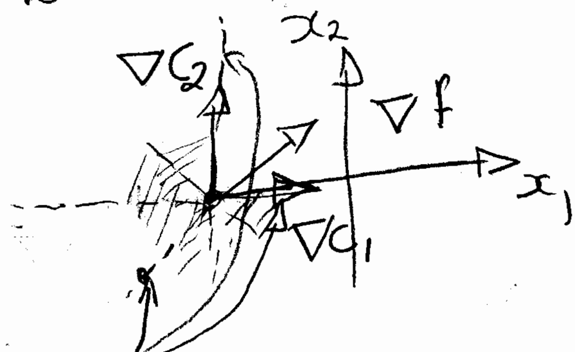
$(x_1, x_2) = \begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$ where both x_1 & x_2 are minimum.

We again have:

$$\nabla g_1(x)^T d \geq 0$$
$$\nabla g_2(x)^T d \geq 0$$
$$\nabla f(x)^T d < 0$$

for a non-minimum direction.

At $\begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$:



d must have positive components

(impossible!)

$\Rightarrow (-\sqrt{2}, 0)^T$ optimal.

Again, set:

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 c_1(x) - \lambda_2 c_2(x)$$

or:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \text{ some } \lambda^* \geq 0.$$

In our example, we have that:

$$\lambda_1^* c_1(x^*) = 0, \quad \lambda_2^* c_2(x^*) = 0.$$

For the solution to the example:

$$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \underline{\text{not zero}}$$

$$\left.\begin{array}{l} \nabla c_1(x^*) = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}, \\ \nabla c_2(x^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array}\right\} \text{ not allowing } d \text{ projections on } -\nabla f(x^*),$$

with $\lambda_1^* = 1/2\sqrt{2}$, $\lambda_2^* = 1$:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

12.2 First-Order Optimality Conditions

Recall the problem:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases}$$

Form the Lagrangian over all constraints:

$$\boxed{\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)}$$

Define the <u>active set</u> at any point x, as:

$$\boxed{\mathcal{A}(x) = \mathcal{E} \cup \{ i \in \mathcal{I} \mid c_i(x) = 0 \}},$$

in other words, $\mathcal{A}(x)$ captures the "boundaries" of the inequality constraints.

We also need to be concerned with degenerate cases, where ∇c_i vanish at the boundaries.

Eg: $c_1(x) = (x_1^2 + x_2^2 - 2)^2 = 0$

gives $\nabla c_1 = 2(x_1^2 + x_2^2 - 2) \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = 0$

on the circle:

$(x_1^2 + x_2^2 - 2) = 0$ where c_1 is active!

LICQ avoids this:

Def 12.1

At any point x, consider its active set $A(x)$. Then, LICQ holds if:

$\{\nabla c_i(x), i \in A(x^*)\}$ is linearly independent.

Clearly then: $\nabla c_i(x) \neq 0$, any i.

First-Order Necessary Conditions

Thm 12.1 Suppose that x^* solves:

$$\min_{x \in \mathbb{R}^n} f(x) \text{ subject to } \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases},$$

and LICQ holds at x^*. Then, we can always find λ_i such that:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0.$$

$c_i(x^*) = 0$, all $i \in \mathcal{E}$, $c_i(x^*) \geq 0$, all $i \in \mathcal{I}$,

$\lambda_i^* c_i(x^*) = 0$, all $i \in \mathcal{E} \cup \mathcal{I}$, $\lambda_i^* \geq 0$, $i \in \mathcal{I}$.

All conditions are known as the Karush-Kuhn-Tucker conditions, or KKT conditions.

Using: $\lambda_i^* c_i(x^*) = 0$, we simplify:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$$

<u>to:</u>

$$0 = \nabla_x \mathcal{L}(x^*, \lambda^*)$$

$$= \nabla f(x^*) - \sum_{\underbrace{i \in A(x^*)}} \lambda_i^* \nabla c_i(x^*)$$

the rest of the terms have $\lambda_i^* = 0$, while these ones have $c_i(x^*) = 0$ to be in $A(x^*)$

<u>From</u> $\lambda_i^* c_i(x^*) = 0$, we have <u>strict complementarity</u> if $\lambda_i^* = 0$ <u>or</u> $c_i(x^*) = 0$ <u>but not both</u> (replace <u>or</u> by <u>exclusive-or</u>)

This means that: $\lambda_i^* > 0$ for $i \in I \cap A(x^*)$

(where $c_i(x^*) = 0$ ...)

When LICQ holds, the solution is unique.

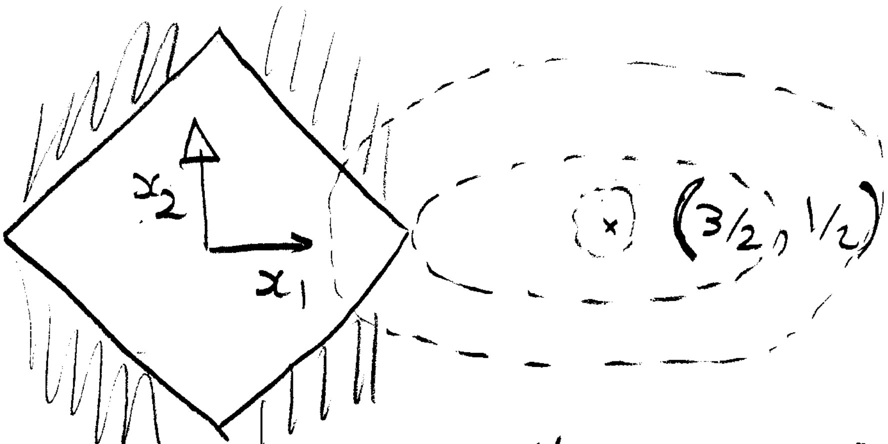
Ex. 12.4

Consider $\min_x \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^4 = f(x)$

such that $\|x\|_1 \leq 1$ or: $|x_1| + |x_2| \leq 1$

or: $\begin{bmatrix} 1 - x_1 - x_2 \\ 1 - x_1 + x_2 \\ 1 + x_1 - x_2 \\ 1 + x_1 + x_2 \end{bmatrix} \geq 0$, which

looks like:



Solution is: $x^* = (1, 0)$.

Plug-in x^* into the $c_i(x)$ to see if $c_i(x^*)=0$:

$$\begin{bmatrix} 1-1-0 \\ 1-1+0 \\ 1+1-0 \\ 1+1+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 2 \end{bmatrix} \begin{matrix} \Rightarrow \begin{cases} c_1(x^*)=0 \\ c_2(x^*)=0. \end{cases} \\ \Rightarrow \begin{cases} \lambda_3^*=0 \\ \lambda_4^*=0 \end{cases} \end{matrix}$$

Now c_1, c_2 <u>are active</u> $\Rightarrow$

$$\nabla f(x^*) = \begin{bmatrix} -1 \\ -1/2 \end{bmatrix} = \lambda_1^* \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_2^* \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad \lambda^* = \begin{bmatrix} 3/4 \\ 1/4 \\ 0 \\ 0 \end{bmatrix} \quad \blacksquare$$

Sensitivity

Note that for $\lambda_i^* = 0$, we have no real constraints! So, take $\lambda_i > 0$, and consider the dependence of f on c_i.

For $-\varepsilon \|\nabla c_i(x^*)\|$ change on c_i:

$$-\varepsilon \|\nabla c_i(x^*)\| = c_i\big(x^*(\varepsilon)\big) - c_i(x^*)$$
$$\approx \big(x^*(\varepsilon) - x^*\big)^T \nabla c_i(x^*)$$

while:

$$0 = c_j\big(x^*(\varepsilon)\big) - c_j(x^*)$$
$$\approx \big(x^*(\varepsilon) - x^*\big)^T \nabla c_j(x^*)$$
$$\text{all } j \in A(x^*),\ j \neq i$$

Also:

$$f(x^*(\varepsilon)) - f(x^*) \simeq \big(x^*(\varepsilon) - x^*\big)^T \nabla f(x^*)$$

But, at x^*: $\nabla f(x^*) = \sum_{j \in A(x^*)} \lambda_j \nabla c_j(x^*)$

Thus:

$$f(x^*(\varepsilon)) - f(x^*) \simeq \sum_{j \in A(x^*)} \lambda_j^* (x^*(\varepsilon) - x^*)^T \nabla c_j(x^*)$$

$$\simeq -\varepsilon \lambda_i^* \|\nabla c_i(x^*)\|$$

from $\textcircled{\Delta}$.

Divide by ε and take $\varepsilon \to 0$ to get:

$$\boxed{\frac{df(x^*(\varepsilon))}{d\varepsilon} = -\lambda_i^* \|\nabla c_i(x^*)\|}$$

... linear dependence.

independent of c_i scaling

($c_i \to 5c_i$ "same")

Strongly-active or Binding:

If $i \in A(x^*)$ and $\lambda_i^* > 0$.

Weakly-active: if $\lambda_i^* = 0$, $i \in A(x^*)$

12.3 Derivation of the First-Order Conditions

* key to understanding <u>all</u> constrained optimization algorithms.

Feasible Sequences

Given a feasible point x^*, a sequence $\{z_k\}_{k=0}^{\infty}$ with $z_k \in \mathbb{R}^n$ is a feasible sequence

<u>if</u>:

(i) $z_k \neq x^*$, all k,

(ii) $\lim_{k \to \infty} z_k = x^*$,

(iii) z_k is feasible for all sufficiently large values of k.

We also let $T(x)$ denote all feasible sequences approaching x.

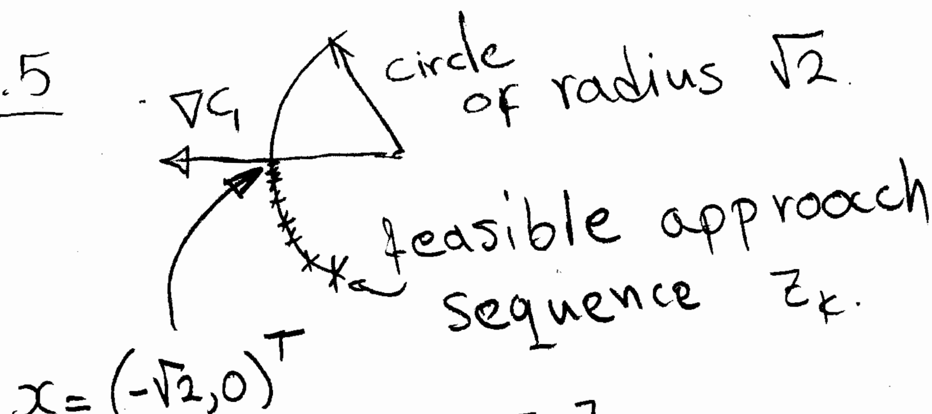
Redefine "local solution" as one for which all feasible sequences satisfy: $f(z_k) \geq f(x)$ for all k sufficiently large.

For any point x, consider:

$$d_k = \frac{z_k - x}{\|z_k - x\|}.$$

Now $\|d_k\| = 1$, and thus all d_k lie on the surface of a multidimensional unit sphere, a <u>compact set</u>. Since $z_k \to x^*$, $d_k \to d$, some direction(s) on this compact set. The d-values are called <u>limiting directions</u>.

<u>Ex 12.5</u>



Let $z_k = \begin{bmatrix} -\sqrt{2 - 1/k^2} \\ -1/k \end{bmatrix}$

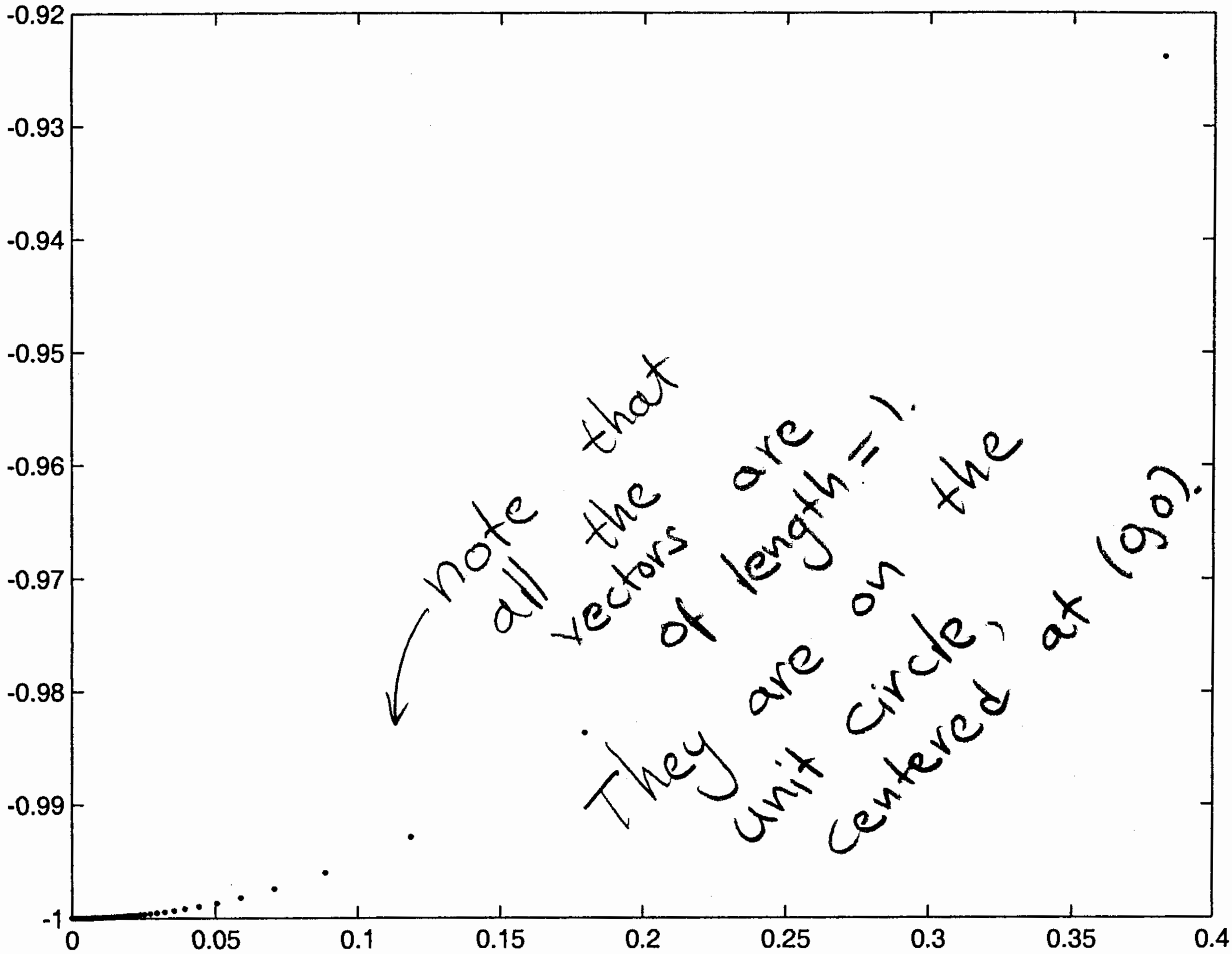
$$d_k = \frac{z_k - x}{\|z_k - x\|} \longrightarrow \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

(see attached)

20/30

```
k=1:100000;
x = -sqrt(2 - 1 ./k.^2);
y = -1 ./ k;
xn = x + sqrt(2);
plot(xn ./ sqrt(xn.^2 + y.^2), y ./ sqrt(xn.^2 + y.^2), '.')
```

Code to generate the limit direction.


-0.92
-0.93
-0.94
-0.95
-0.96
-0.97
-0.98
-0.99
-1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
note that all the vectors are of length = 1.
They are on the unit circle, centered at (0,0).

Similarly, for

$$z_k = \begin{bmatrix} -\sqrt{2 - 1/k^2} \\ 1/k \end{bmatrix},$$ we have:

$$d_k \rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

<u>Any</u> feasible sequence approaching $(-\sqrt{2}, 0)^T$ will have two limiting directions: $(0,1)^T$ and $(0,-1)^T$.

If the objective function increases in its first-order approx along any $\forall \{z_k\} \in T(x)$ then x <u>must not be optimal</u>.

Thm 12.2 If x^* is a local solution of 12.1, then all feasible sequences $\{z_k\}$ in $T(x^*)$ must satisfy:

$$\nabla f(x^*)^T d \geqslant 0,$$

where d is any limiting direction of $\{z_k\}$.

Proof (by contradiction).

Assume $\{z_k\}$ is feasible with $\nabla f(x^*)^T d < 0$, for some limiting direction d. Then, for subsequence: $z_k \to x^*$, by Taylor:

$$f(z_k) = f(x^*) + (z_k - x^*)^T \nabla f(x^*) + o(\|z_k - x^*\|)$$

$$= f(x^*) + \underbrace{\|z_k - x^*\| d^T \nabla f(x^*)}_{<0} + o(\|z_k - x^*\|)$$

Since $d^T \nabla f(x^*) < 0$, and the 2nd term dominates the third as $k \to \infty$:

$$f(x^*) > f(z_k), \text{ impossible!}$$

(correct your book).

Characterizing Limiting Directions: Constraint Qualifications

Define: $\nabla c_i^* = \nabla c_i(x^*)$,

$$A^T = \left[\nabla c_i^*\right]_{i \in A(x^*)},$$

$$\nabla f^* = \nabla f(x^*)$$

and recall:

$$A(x^*) = \mathcal{E} \cup \{ i \in \mathcal{I} \mid c_i(x) = 0 \}$$

where constraints are still active.

Lemma 12.3

(i) If d is a limiting direction of a feasible sequence, then:

$$\left.\begin{aligned} d^T \nabla c_i^* &= 0, \text{ all } i \in \mathcal{E} \\ d^T \nabla c_i^* &\geq 0, \text{ all } i \in A(x^*) \cap \mathcal{I} \end{aligned}\right\} (*)$$

(ii) If $(*)$ holds with $\|d\| = 1$ and LICQ holds, then d is a limiting direction of some feasible sequence.

Proof of (i):

Let $\{z_k\}$ be a feasible sequence, and let d be its limiting direction. Then either z_k or a subsequence of it must satisfy:

$$\lim_{k\to\infty} \frac{z_k - x^*}{\|z_k - x^*\|} = d$$

which gives:

$$z_k = x^* + \|z_k - x^*\| d + o(\|z_k - x^*\|)$$

For active constraints:

directional derivative in d direction

$$0 = \frac{1}{\|z_k - x^*\|} \overbrace{c_i(z_k)}^{=0}$$

$$= \frac{1}{\|z_k - x^*\|} \Big[\overbrace{c_i(x^*)}^{=0} + \|z_k - x^*\| \nabla c_i^T d$$

$$+ o(\|z_k - x^*\|) \Big]$$

$$= \nabla c_i^T d + \frac{o(\|z_k - x^*\|)}{\|z_k - x^*\|} \longrightarrow \nabla c_i^T d \quad \text{as } k \to \infty$$

For active constraints, replace "$0 =$" by "$0 \leq$".

Def 12.4

Given x^* and the active constraint set $\mathcal{A}(x^*)$, define F_1 by:

$$F_1 = \left\{ \alpha d \;\middle|\; \alpha > 0 \text{ with: } \begin{array}{l} d^T \nabla c_i^* = 0, \ i \in \mathcal{E} \\ d^T \nabla c_i^* \geq 0, \ i \in \mathcal{A}(x^*) \cap \mathcal{I} \end{array} \right\}$$

Lemma 12.4

There is no direction $d \in F_1$ for which:

$d^T \nabla f^* < 0$, <u>iff</u> $\exists \lambda \in \mathbb{R}^n$ with:

$$\nabla f^* = \sum_{i \in \mathcal{A}(x^*)} \lambda_i \nabla c_i^* = A(x^*)^T \lambda,$$

$$\lambda_i \geq 0 \quad \text{for } i \in \mathcal{A}(x^*) \cap \mathcal{I}.$$

<u>Proof:</u> omitted.

Thm 12.1: Proof outline.

→ From $(***)$:

$$\nabla f^* = \sum_{i \in A(x^*)} \lambda_i \nabla c_i^*,$$

conclude that: $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$,

and $\lambda_i \geq 0$ for active constraints,

or $\lambda_i^* = 0$ in the interior.

12.4 Second-Order Conditions

Thm 12.5 Suppose that:

* x^* is a local solution and LICQ is satisfied,
* Let λ^* be a Lagrange multiplier vector satisfying the KKT conditions,
* Define $w \in F_2(\lambda^*)$ by:

$$w \in F_2(\lambda^*) \Longleftrightarrow \begin{cases} \nabla c_i(x^*)^T w = 0, & \text{all } i \in \mathcal{E}, \\ \nabla c_i(x^*)^T w = 0, & \text{all } i \in A(x^*) \cap \mathcal{I} \text{ with } \lambda_i^* > 0, \\ \nabla c_i(x^*)^T w \geq 0, & \text{all } i \in A(x^*) \cap \mathcal{I}, \text{ with } \lambda_i^* = 0 \end{cases}$$

Then:

$$\omega^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*)\omega \geq 0,$$

for all $\omega \in F_2(\lambda^*)$.

Thm 12.5 gives necessary conditions.

Thm 12.6 gives sufficient conditions for a strict local minimum:

Thm 12.6 Suppose that $x^* \in \mathbb{R}^n$ is feasible, λ^* can be found that satisfies the KKT conditions. Also, suppose:

$$\omega^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*)\omega > 0, \quad \omega \in F_2(\lambda^*), \omega \neq 0$$

Then x^* is a strict local solution.

NB: LICQ is gone and $>$ replace $\geq$.

Note that thms 12.5, 12.6 cannot directly lead to a test of positive definiteness or semi-definiteness because ws are constrained to F_2.

We can correct this by considering "projected Hessians":

<u>Basic idea:</u>

1. Construct Z from the null space of A
2. Check:

$$Z^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*) Z$$

<u>for:</u>

2(a) positive semi-definiteness (thm 12.5), or

2(b) positive definiteness (thm 12.6).

For constructing the nullspace of A, see Strang's discussion on the SVD.

Now, if $\nabla_{xx}\mathcal{L}(x^*, \lambda^*)$ is positive definite/semi-definite, then thms 12.5, 12.6 apply. But this is overly restrictive.

We can also reduce our requirements from LICQ to MFCQ:

Def 12.5: For $x^* \in A(x^*)$, MFCQ holds if we can find $w \in \mathbb{R}^n$ with:

* $\nabla c_i(x^*)^T w > 0, \quad i \in A(x^*) \cap \mathcal{I}$
* $\nabla c_i(x^*)^T w = 0, \quad i \in \mathcal{E}$.
* $\{\nabla c_i(x^*), i \in \mathcal{E}\}$ is linearly independent.