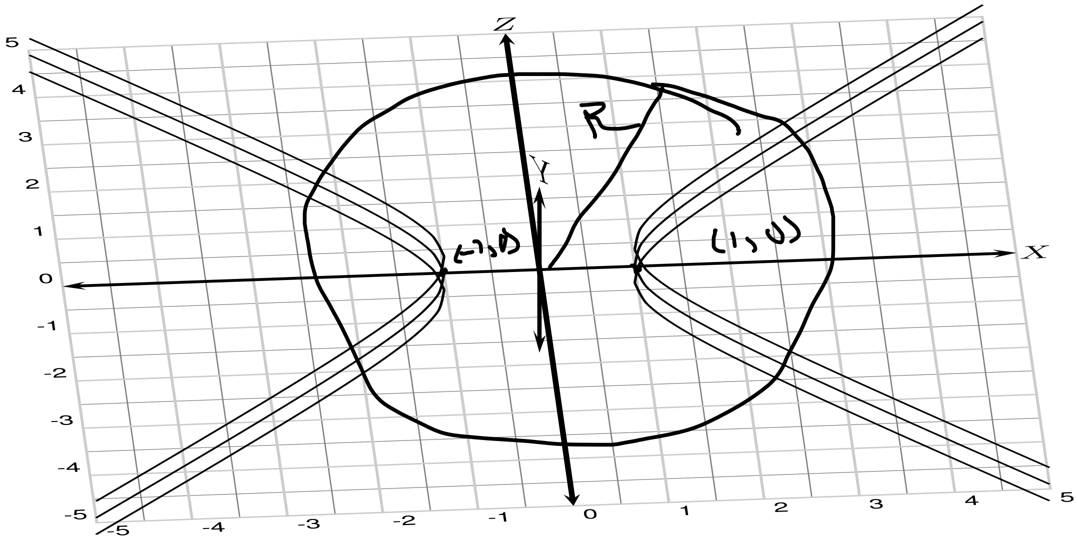
$\min \ x^2+y^2+z^2$

$x^2 - z^2 = 1$ $\quad x = \sqrt{1+z^2}$

30 Lecture 30

[30.1] **Application of SOSC** Find a point (x, y, z) on the surface $x^2 - z^2 = 1$ which is nearest to the origin.

Answer.



a) Note that, from the figure, it is visible that the points are $\pm e_1$. But let us argue it, assuming that we do not have the picture.

b) Sometimes constrained optimization problems can be converted to unconstrained optimization problems. Here the problem is

$1+z^2$

$$\begin{array}{ll} \min & x^2+y^2+z^2 \\ \text{s.t.} & x^2-z^2=1 \end{array} \quad \equiv \quad \begin{array}{ll} \min & y^2+2z^2+1 \\ \text{s.t.} & y, z \in \mathbb{R} \end{array}.$$

$f = y^2+2z^2+1$

$\nabla f = \begin{bmatrix} 2y \\ 4z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

c) The rhs is an unconstrained problem in two variables. We solve that.

d) Critical points: We have only one critical point $\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$H_f(a)$ $\quad$ crit $\to \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad$ $H(a)$

e) Time to use SOSC: The Hessian $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ is pd. So by SOSC, it is a strict local minimum.

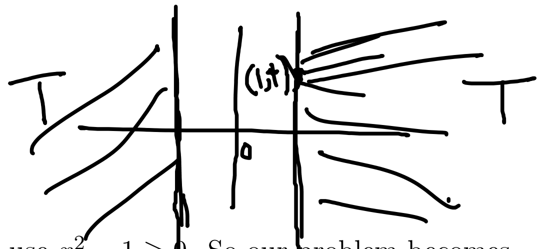
$\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$

f) As the function $y^2 + 2z^2 + 1$ is always ≥ 1, we see that it is a strict absolute minimum.

g) For the original problem: we have $x^2 = 1 + z^2 = 1$. Hence $x = \pm 1$, $y = 0$, $z = 0$ are strict absolute minimums.

h) Can we conclude them to be global minimum in some other way also? Of course. Take a big r and consider the part of the surface lying inside $\overline{B_r(0)}$. That becomes a closed and bounded set, that is, a compact set. As $f(x, y, z) = \|(x, y, z)\|^2$ is a continuous function, it will attain its minimum. This point of minimum, must be a global minimum, as points outside the ball have larger distance from origin. It will also be a critical point. Since $f(e_1)$ is the same as $f(-e_1)$, they both must be global minimums. (Otherwise, we should have another critical point.)

$\nabla f = \begin{bmatrix} 4x \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $z^2 = x^2 - 1$



Alternate. A) Alternately, we can use $z^2 = x^2 - 1$. Then we must use $x^2 - 1 \geq 0$. So our problem becomes

$$\begin{array}{ll} \min & 2x^2 + y^2 - 1 \\ \text{s.t.} & x^2 \geq 1,\ y \in \mathbb{R}. \end{array}$$

min $2x^2 + y^2 - 1$

$x^2 \geq 1$, $y \in \mathbb{R}$

B) FONC: We have no critical points in the interior. ✓

C) FONC: For a point $a = (1, t)$, where $t \neq 0$, we have $\nabla f(a) = \begin{bmatrix} 4 \\ 2t \end{bmatrix}$, feasible directions are $\begin{bmatrix} \geq 0 \\ * \end{bmatrix}$ and $\nabla f(a)^t \begin{bmatrix} 0 \\ -1 \end{bmatrix} < 0$. So this point is not a local minimum. Similarly, $(-1, t)$, $t \neq 0$, is not a point of local minimum.

$\nabla f(a) = \begin{bmatrix} 4x \\ 2y \end{bmatrix} = \begin{bmatrix} 4 \\ 2t \end{bmatrix}$ $\begin{bmatrix} \geq 0 \\ * \end{bmatrix}$, $t \neq 0$

$\langle \nabla f(a), \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix} \rangle = \pm 2t$

$(1,t)$, $t \neq 0$ do not satisfy FONC

$\left[\begin{array}{l} \nabla f(a)^t d \geq 0 \quad \forall d \in \bar{D}(a) \\ \nabla f(a)^t d \geq 0, \ D(a) \ ✓ \end{array} \right.$

D) FONC: The point $a = (1, 0)$ satisfies FONC for being a local minimum as

$$\nabla f(a)^t d = \begin{bmatrix} 4 & 0 \end{bmatrix} \begin{bmatrix} \geq 0 \\ * \end{bmatrix} \geq 0, \qquad \forall d \in D(a). \quad \text{(Remember: checking } D(a) \text{ is enough.)}$$

$\nabla f = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\begin{bmatrix} \geq 0 \\ * \end{bmatrix}$ ≥ 0

$d = \begin{bmatrix} 0 \\ * \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix}$

E) SONC: The directions $d \in D(a)$ such that $\nabla f(a)^t d = 0$ are $\alpha \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix}$. As

$$e_2^t H(a) e_2 = e_2^t \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} e_2 = 2 \geq 0,$$

we see that the point a satisfies SONC for being a local minimum.

$d^t H(a) d \geq 0$

$\alpha \begin{bmatrix} 0 & \pm 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \alpha \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix}$

$= 2\alpha^2 \geq 0.$

F) Similarly, $(-1, 0)$ satisfies FONC and SONC.

G) Since $(1, 0)$ is not a point of interior, we cannot use SOSC. However, we can conclude the minimality sometimes using some other observations.

Since the function is large outside a ball, we see that f must have an absolute minimum. At these points

FONC and SONC must be satisfied. We have only two such points $(\pm 1, 0)$. The value of the function is the same at these points. So both these points are global minimums.

Some exercises

[30.2] **Exercises(E)** Let $f \in \mathcal{C}^2(T)$ and $a \in T$ be a point at which FONC and SONC holds. Suppose that a is not a point of interior but the Hessian $H(a)$ is positive definite. Show that a may not be a point of minimum.

[30.3] **Exercise(M)** `(Why is it happening?)`

a) I have a convex cube in $\mathbb{R}^3$. Suppose that I have a linear function which takes equal values at two diametrically opposite vertices. Must that function be a constant?

b) I have a convex cube in $\mathbb{R}^3$. Suppose that I have a linear function which is minimized at two diametrically opposite vertices. Must that function be a constant?

[30.4] **NoPen**

a) Let $f \in \mathcal{C}^2(T)$ and $a \in T^o$ be a critical point. In SOSC item b), it says that if $H(x)$ is psd in a neighborhood $B_\delta(a)$, then a is local minimum. Can we go with just that $H(a)$ is psd?

b) Give an example of a closed convex feasible set and a point for which $D(a)$ is not closed.

c) Consider minimizing $c^t x$, $c \neq 0$ over a set T. Can we have an optimum at an interior point?

d) Let $A \in M_n(\mathbb{R})$ and $f : \mathbb{R}^n \to \mathbb{R}$ be defined as $f(x) = x^t A x$. What is $\bigtriangledown f(x)$? What is $H(x)$?

e) Let $A \in M_3(\mathbb{R})$ and $b \in \mathbb{R}^3$. Must the system $Ax = b$ have at least one solution? What if A is psd? What if A is pd?

[30.5] **Exercise(E)** Let T be convex and $a \in T$. Is $D(a)$ necessarily a convex cone?

[30.6] **Exercise(E)** Let $A \in M_n(\mathbb{R})$, $b \in \mathbb{R}^n$, and consider $f : \mathbb{R}^n \to \mathbb{R}$ defined as $f(x) = x^t A x + b^t x$.

a) What is $\bigtriangledown f(x)$? What is $H(x)$?

b) Suppose that A is psd. Is it necessary that we should have at least one critical point?

c) Suppose that A is psd. Suppose that a is a critical point. Can it be a saddle point/ local maximum/ local minimum?

d) Let A be positive definite. Is it necessary that we should have at least one critical point? Check whether they are minimums or maximums or saddle points.

[30.7] **Practice** Find local optimums and saddle points of $f = 2x_1x_2x_3 - 4x_1x_3 - 2x_2x_3 + x_1^2 + x_2^2 + x_3^2 - 2x_1 - 4x_2 + 4x_3$.

[30.8] **Practice** Find the local optimums and saddle points of $f(x, y) = 5x^3 + 4xy + x + y^2$.

[30.9] **Exercise(E)** Let $v_1, \ldots, v_p \in \mathbb{R}^n$ be distinct and define $f : \mathbb{R}^n \to \mathbb{R}$ as $f(x) = \sum\limits_{i=1}^{p} \|x - v_i\|^2$. Optimize it.

[30.10] **Exercise(E)** Let $v_1 < \cdots < v_p$ be some real numbers and define $f : \mathbb{R} \to \mathbb{R}$ as $f(x) = \sum_{i=1}^{p} |x - v_i|$. Optimize it.

[30.11] **Exercise(E)** Consider $\begin{array}{ll} \max & cx + dy \\ \text{s.t.} & x + y \le 1,\ x, y \ge 0, \end{array}$ where $c > d \ge 0$. Use FONC to show that the unique maximum solution is $(1,0)^t$. Use graphical method to give an alternate solution.

Constrained optimization

$T = \left\{ \begin{array}{l} g_1(x) \geq 0 \\ g_m(x) \geq 0 \\ h_1(x) = 0 \\ h_p(x) = 0 \end{array} \right\}$

[30.12] **Definitions**

a) Suppose that the set T is defined using some $\mathcal{C}^1$ functions,

$$T = \{x \in \mathbb{R}^n \mid g_i(x) \ge 0, i = 1, \ldots, m, \quad h_j(x) = 0, j = 1, \ldots, p\}.$$

Notice here 'the inequality constraints are in '$g_i(x) \ge 0$' form.

b) A constraint is said to be an ACTIVE CONSTRAINT at a point $a \in T$, if it satisfies equality in the constraint.

c) Since all h_j are active at each feasible points, let us denote by $\boldsymbol{A(a)}$ the set

$a \in T$

$$A(a) := \{i \mid g_i \text{ is active at } a\}.$$

$A(a) = \{ i \mid g_i(a) = 0,\ i = 1, \ldots, m \}$

d) The LINEARIZING CONE $\boldsymbol{\mathcal{D}(a)}$ of T at a is defined as

$$\mathcal{D}(a) := \Big\{ d \in \mathbb{R}^n \mid \langle \bigtriangledown g_i(a), d \rangle \ge 0, \forall i \in A(a), \quad \langle \bigtriangledown h_j(a), d \rangle = 0, \forall j \Big\}.$$

e) It is a nonempty closed convex cone.‼ Loosely saying, it gives us the directions along which we can move a little bit and still stay inside the feasible set, that is, $\overline{D}(a)$. (We will show this later.)

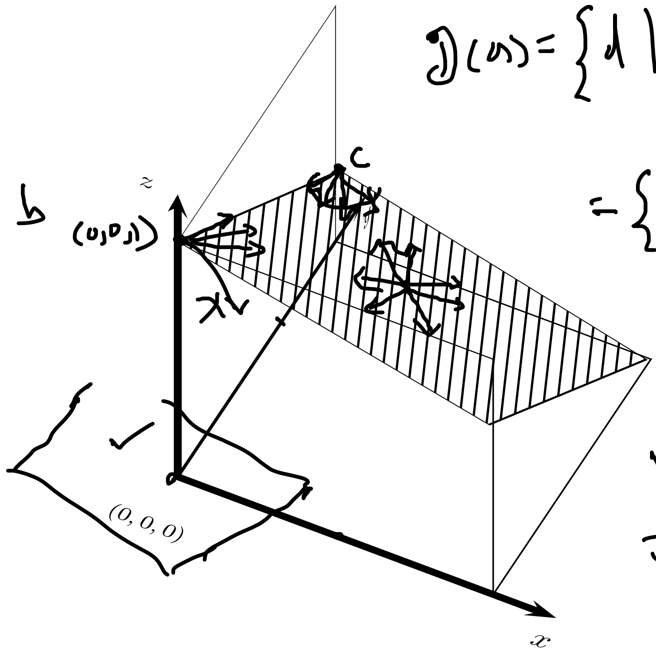
$g_i(a + td) \geq 0$

$\mathcal{D}(a) = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \nabla g_i(a)^t d \geq 0 & i \in A(a) \\ \nabla h_j(a)^t d = 0 & j = 1, 2, \ldots, p \end{array} \right\}$

[30.13] Example Let T be the intersection of the closed unit cube with corners $(0,0,0)$ and $(1,1,1)$ and the hyperplane $x+2y+3z=3$. That is,

$$T = \{(x,y,z)^t \in \mathbb{R}^3 \mid g_1 \equiv 1-x \ge 0, \quad g_2 \equiv 1-y \ge 0, \quad g_3 \equiv 1-z \ge 0,$$
$$g_4 \equiv x \ge 0, \quad g_5 \equiv y \ge 0, \quad g_6 \equiv z \ge 0, \quad h_1 \equiv x+2y+3z-3 = 0\}.$$

$0 \le x \le 1$
y
z
$1-x \ge 0$ ✗
g_1 $1-y \ge 0$ ✗
g_2 $1-z \ge 0$ ✓
g_3 $x \ge 0$ ✓
$y \ge 0$ ✓
$z \ge 0$ ✗
g_6
h_1: $x+2y+3z=3$



$\mathcal{D}(a) = \{d \mid \cancel{\nabla g_i^t d \ge 0 \ \forall i \in A(a)}$
$\nabla h_1 d = 0\}$
$= \{d \mid d_1 + 2d_2 + 3d_3 = 0\}$
$= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^{\perp}$
$\parallel$
$\mathcal{D}(a) = \bar{\mathcal{D}}(a)$

$\mathcal{D}(b) = \{d \mid d_3 \le 0, d_1 \ge 0, d_2 \ge 0, d_1 + 2d_2 + 3d_3 = 0\}$
$\parallel$
$\mathcal{D}(b) = \bar{\mathcal{D}}(b)$

∘ Let us take $a = (.5, .5, .5)$. Then $A(a) = \emptyset$. We have

$$\mathcal{D}(a) = \{d \mid \nabla h_1^t d = 0\} = \{d \mid \begin{bmatrix}1 & 2 & 3\end{bmatrix} d = 0\},$$

a plane parallel to our h_1-plane passing through the origin.

∘ If we take $b = (0,0,1)$, then $A(b) = \{3,4,5\}$. We have

$$\begin{aligned} \mathcal{D}(b) &= \{d \mid \nabla g_3^t d \ge 0, \quad \nabla g_4^t d \ge 0, \quad \nabla g_5^t d \ge 0, \quad \nabla h_1^t d = 0\} \\ &= \{d \mid d_3 \le 0, \quad d_1 \ge 0, \quad d_2 \ge 0, \quad d_1 + 2d_2 + 3d_3 = 0\}. \end{aligned}$$

Notice that, if we stand at $b = (0,0,1)$ and look at our feasible region, then these are the feasible directions available at this point.

[30.14] Practice Consider $T = \{x \in \mathbb{R}^2 \mid x \geq 0, \quad y \geq 0, \quad 1 - x^2 - y^2 \geq 0\}$. Take $a = (1,0)$ and $b = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Compute $D(a), D(b), \mathcal{D}(a)$ and $\mathcal{D}(b)$.

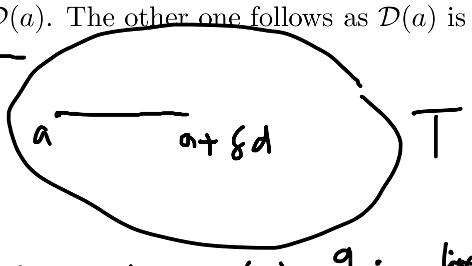
[30.15] Proposition ($\mathcal{D}(a)$ contains $\overline{D}(a)$) Let

$$T = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, m, \quad h_j(x) = 0, j = 1, \ldots, p\}$$

and $a \in T$. Then $\overline{D}(a) \subseteq \mathcal{D}(a)$. The inclusion can be strict.

Proof. a) First we show $D(a) \subseteq \mathcal{D}(a)$. The other one follows as $\mathcal{D}(a)$ is closed.

Let $d \in D(a)$.



To show $d \in \mathcal{D}(a) = \{d \mid \nabla g_i(a)^t d \geq 0 \quad i \in A(a), \quad \nabla h_j(a)^t d = 0 \ \forall j\}$ ✓

Let g_i be active at a.

So $g_i(a) = 0$.

$$D_d g_i(a) = \lim_{t \to 0+} \frac{\overbrace{g_i(a+td)}^{\geq 0} - \overbrace{g_i(a)}^{=0}}{t} \geq 0$$ ✓

$$\nabla h_j(a)^t d = D_d h_j(a) = \lim_{t \to 0+} \frac{\overbrace{h_j(a+td)}^{0} - \overbrace{h_j(a)}^{0}}{t} = 0$$

b) Let $d \in D(a)$. So there exists $\delta > 0$ such that $[a, a + \delta d] \subseteq T$.

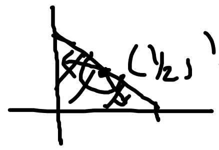
$g_1 = (1-x-y)^3 \geq 0$

$g_2 = x \geq 0$, $g_3 \geq y \geq 0$

c) If $g_i(a) > 0$, we do not bother.

d) If $g_i(a) = 0$, as $a + td \in T$ for all $0 \leq t \leq \delta$, we see that $g_i(a + td) \geq 0$. So $\nabla g_i(a)^t d \geq 0$.

e) Similarly, $\nabla h_j(a)^t d = 0$.



f) So $d \in \mathcal{D}(a)$. As $\mathcal{D}(a)$ is a closed set, $\overline{D}(a) \subseteq \mathcal{D}(a)$.

g) To see that the inclusion can be strict, let $a = (\frac{1}{2}, \frac{1}{2}) \in T = \{(x,y) \mid g \equiv (1 - x - y)^3 \geq 0, x, y \geq 0\}$. The only constraint active at a is g. Further, $\nabla g(a) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, so that $\mathcal{D}(a) = \mathbb{R}^2$. We see that

not used g_2, g_3

$$\begin{aligned} D(a) &= \left\{d \mid g(a + \theta d) \geq 0, \forall \theta \in (0, \delta), \text{ for some } \delta > 0\right\} = \left\{d \mid \left(1 - .5 - \theta d_1 - .5 - \theta d_2\right)^3 \geq 0\right\} \\ &= \left\{(d_1, d_2) \mid \left(\theta d_1 + \theta d_2\right)^3 \leq 0\right\} = \left\{(d_1, d_2) \mid d_1 + d_2 \leq 0\right\}. \end{aligned}$$

Hence $\overline{D}(a) = D(a) \subsetneq \mathbb{R}^2 = \mathcal{D}(a)$. [Alternately, observe that, the constraint g is the same as $x + y \leq 1$. So $\overline{D}(a) = \{d \mid d_1 + d_2 \leq 0\}$.] ■

$a = (\frac{1}{2}, \frac{1}{2})$ $\quad \mathcal{D}(a) = \{d \mid \nabla g_1^t(a) d \geq 0\} = \mathbb{R}^2$

$D(a) = \{d \mid a + td \in T, \ 0 \leq t \leq \delta_d$ for some $\delta_d\} = \{d \mid d_1 + d_2 \leq 0\}$

$\subsetneq \mathbb{R}^2$