[31.8] **Theorem** Let a be a feasible point of (P2). TFAE.

a) The set $Z(a) = \emptyset$.

b) There exist $\lambda_i \geq 0$ and $w_j \in \mathbb{R}$, such that $\nabla L(a, \lambda, w) = 0$ and $\lambda_i g_i(a) = 0$ for each i.

Proof. a)⇒b).

$$Z(a) = \left\{ d \;\middle|\; \begin{array}{ll} \nabla g_i^t d \geq 0 & \forall \text{ active } g_i \\ \nabla h_j^t d = 0 & \forall j = 1, \dots, p \\ \nabla f(a)^t d < 0 & \end{array} \right\} = \emptyset$$

This means $d \in D(a) \Rightarrow \nabla f(a)^t d \geq 0$.

Assume $g_1, \dots, g_k$ are active.

$B^t d \geq 0 \Rightarrow c^t d \geq 0$

$$\underbrace{\begin{bmatrix} \nabla g_1^t(a) \\ \vdots \\ \nabla g_k^t \\ \hline \nabla h_1^t \\ \vdots \\ \nabla h_p^t \\ \hline -\nabla h_1^t \\ \vdots \\ -\nabla h_p^t \end{bmatrix}}_{B^t} d \geq 0 \Rightarrow \nabla f^t d \geq 0$$

Farkas's Lemma $\exists\ y \geq 0$ s.t. $\nabla f^t = y^T B^t$ or $\nabla f = By$

$$\nabla f = \begin{bmatrix} \nabla g_1 & \cdots & \nabla g_k & \nabla h_1 & \cdots & \nabla h_p & -\nabla h_1 & \cdots & -\nabla h_p \end{bmatrix} \begin{bmatrix} y_1 \\ y_k \\ y_{k+1} \\ \\ y_{k+2p} \end{bmatrix}$$

$$= y_1 \nabla g_1 + \dots + y_k \nabla g_k + (y_{k+1} - y_{k+p+1}) \nabla h_1 + \dots$$

$$+ (y_{k+p} - y_{k+2p}) \nabla h_p$$

$\lambda_1, \dots, \lambda_k \geq 0$
$w_i \in \mathbb{R}$.

$$= \lambda_1 \nabla g_1 + \quad + \lambda_k \nabla g_k + w_1 \nabla h_1 + \dots + w_p \nabla h_p$$

$$= \lambda_1 \nabla g_1 + \dots + \lambda_k \nabla g_k + \underset{\lambda_{k+1}}{0} \nabla g_{k+1} + \dots + \underset{\lambda_m}{0} \nabla g_m + \qquad \text{is}$$

$\exists \lambda_i \geq 0$
$w_j \in \mathbb{R}$

$$\nabla f = \sum \lambda_i \nabla g_i + \sum w_j \nabla h_j \quad \underline{\text{where}} \quad \lambda_i g_i(a) = 0 \ \forall i$$

Let $Z(a) = \emptyset$. So for each $d \in \mathcal{D}(a)$ we have $\triangledown f(a)^t d \geq 0$. For simplicity, assume that $g_1, \ldots, g_k$ are active at a. A vector $d \in \mathcal{D}(a)$ is nothing but a vector that satisfies $B^t d \geq 0$, where

$$B^t = \begin{bmatrix} \triangledown g_1(a)^t \\ \vdots \\ \triangledown g_k^t(a) \\ \hline \triangledown h_1(a)^t \\ \vdots \\ \triangledown h_p^t(a) \\ \hline -\triangledown h_1(a)^t \\ \vdots \\ -\triangledown h_p(a)^t \end{bmatrix}, \qquad \text{so that } B = \left[\triangledown g_1 \ \cdots \ \triangledown g_k \,\middle|\, \triangledown h_1 \ \cdots \ \triangledown h_p \,\middle|\, -\triangledown h_1 \ \cdots \ -\triangledown h_p\right].$$

Thus

$$B^t d \geq 0 \quad \Rightarrow \quad \triangledown f(a)^t d \geq 0.$$

By Farka's lemma, $\exists y \geq 0$ such that $\triangledown f(a)^t = y^t B^t$, that is,

$$\begin{aligned} \triangledown f(a) &= By \\ &= y_1 \triangledown g_1(a) + \cdots + y_k \triangledown g_k(a) + y_{k+1} \triangledown h_1(a) + \cdots + y_{k+p} \triangledown h_p(a) \\ &\quad -y_{k+p+1} \triangledown h_1(a) - \cdots - y_{k+2p} \triangledown h_p \\ &= y_1 \triangledown g_1(a) + \cdots + y_k \triangledown g_k(a) + (y_{k+1} - y_{k+p+1}) \triangledown h_1(a) + \cdots + (y_{k+p} - y_{k+2p}) \triangledown h_p(a) \\ &= \lambda_1 \triangledown g_1(a) + \cdots + \lambda_k \triangledown g_k(a) + w_1 \triangledown h_1(a) + \cdots + w_k \triangledown h_p(a) \qquad \begin{array}{l}(\text{put } \lambda_i = y_i,\ i = 1, \ldots, k \\ \text{and } w_j = y_{k+j} - y_{k+p+j}, \\ j = 1, \ldots, p)\end{array} \\ &= \lambda_1 \triangledown g_1(a) + \cdots + \lambda_k \triangledown g_k(a) + 0 \triangledown g_{k+1}(a) + \cdots + 0 \triangledown g_m(a) \\ &\quad +w_1 \triangledown h_1(a) + \cdots + w_k \triangledown h_p(a). \end{aligned}$$

Notice that, by construction $\lambda_i \geq 0$ for $i = 1, \ldots, m$ and $w_j \in \mathbb{R}$ for $j = 1, \ldots, p$. Again notice that $\lambda_i g_i(a) = 0$ for each $i = 1, \ldots, m$ due to construction. Indeed, if $g_i(a) = 0$, then $\lambda_i g_i(a) = 0$ and if $g_i(a) > 0$, then $i \notin A(a)$ and $\lambda_i = 0$ (by our choice) implying that $\lambda_i g_i(a) = 0$.

b)$\Rightarrow$a). $\nabla f^t d \geq 0$ for each $d \in \mathcal{D}(a)$

$$\| \quad \{d \mid \nabla g_i^t d \geq 0 \text{ for all active const} \quad \nabla h_j^t d = 0 \text{ for all } h_j\}$$

g_i inactive $\Rightarrow g_i(a) > 0$ given $\lambda_i g_i(a) = 0$ so $\lambda_i = 0$ for inactive

b) $\Rightarrow \exists \lambda_i \geq 0, w_j \in \mathbb{R}$ s.t.

$$\begin{aligned} \nabla f^t d &= \sum \lambda_i \nabla g_i^t d + \sum w_j \nabla h_j^t d \\ &= \sum_{\text{active}} \lambda_i \nabla g_i^t d + \sum_{\text{inactive}} \lambda_i \nabla g_i^t d + \sum w_j \nabla h_j^t d \\ &= \geq 0 + 0 + 0 \\ &\geq 0 \end{aligned}$$

Assume that b) holds. We want to show that $Z(a) = \emptyset$. That is, $\{d \mid d \in \mathcal{D}(a),\ D_d f(a) < 0\} = \emptyset$.

For that, let $d \in \mathcal{D}(a)$. By definition, for each $i \in A(a)$, we have $\bigtriangledown g_i^t d \geq 0$ and for each j we have $\bigtriangledown h_j^t d = 0$.

As $\lambda_i g_i(a) = 0$ holds for each i by the hypothesis of b), we see that $\lambda_i = 0$ for each $i \notin A(a)$.

Hence, from the hypothesis of b), we get that

$$\begin{array}{rcccccc} \bigtriangledown f(a)^t d & = & \sum\limits_{i \in A(a)} \lambda_i \bigtriangledown g_i(a)^t d & + & \sum\limits_{i \notin A(a)} \lambda_i \bigtriangledown g_i(a)^t d & + & \sum_j w_j \bigtriangledown h_j(a)^t d \\ & = & \text{nonnegative (as } d \in \mathcal{D}(a)) & + & 0 \text{ (as } \lambda_i = 0 \text{ here)} & + & 0 \text{ (as } d \in \mathcal{D}(a)) \\ & \geq & 0 & & & & . \end{array}$$

Thus $Z(a) = \emptyset$. ∎

32 Lecture 32

Kuhn-Tucker points

[32.1] **Definition** Consider the problem (P2). A point a is a Kuhn-Tucker point (KT point) if it satisfies the following conditions.

$$\text{KT conditions} \quad \begin{array}{ll} i) & a \in T \\ ii) & \exists \lambda_i \geq 0, w_j, \text{ such that } \bigtriangledown L(a, \lambda, w) = 0 \\ iii) & \lambda_i g_i = 0, \forall i. \end{array}$$

[32.2] **Remark** In view of [31.8], these are the points for which $Z(a) = \emptyset$, that is, the directional derivative of f is nonnegative along each direction in the linearizing cone. So these are possible local minimums.

$\mathbb{R}^{\circ}_{+}$
H°_{+}
Z_{+}

[32.3] **Example** Find all KT points for the problem

$$\begin{array}{ll} \min & x - y + z \\ \hline \text{s.t.} & \begin{array}{l} g_1 \equiv 1 - x \geq 0, g_2 \equiv 1 - y \geq 0, g_3 \equiv 1 - z \geq 0, g_4 \equiv x \geq 0, \\ g_5 \equiv y \geq 0, g_6 \equiv z \geq 0, h_1 \equiv x + 2y + 3z - 3 = 0 \end{array} \end{array}$$

in two different ways.

Answer. Note that $f'(a) = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$ at any point a. There are two ways to find out the KT points.

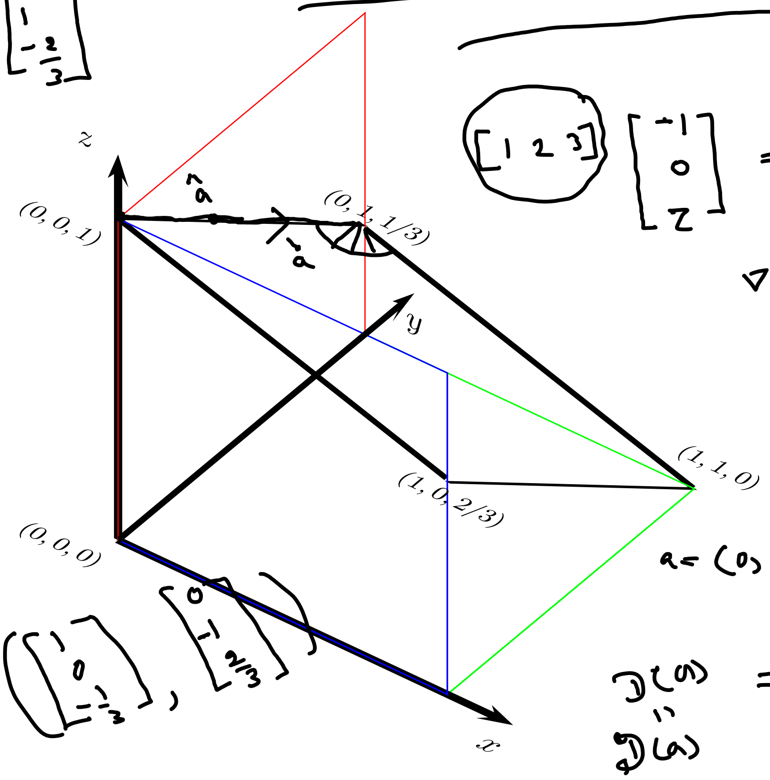
A)`(First (geometric) way.)` Use the fact that $Z(a) = \emptyset$.

Let us first visualize the feasible region. (If we can visualize, then we would see how we are selecting the directions. But our arguments can be verified mechanically.)

$\hat{d} = \begin{bmatrix}0\\1\\-\frac{2}{3}\end{bmatrix}$ $\quad \nabla f^t \hat{d}(\hat{a}) = -5/3 < 0 \quad Z(\hat{a}) \neq \emptyset$

$g_1 \equiv 1-x \geq 0 \quad g_4 \equiv x \geq 0$
$g_2 \equiv 1-y \geq 0 \quad g_5 \equiv y \geq 0$
$g_3 \equiv 1-z \geq 0 \quad g_6 \equiv z \geq 0$

$h \equiv x+2y+3z-3 = 0$



$\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}-1\\0\\z\end{bmatrix} = 0 \Rightarrow z = 1/3. \quad d = \begin{bmatrix}-1\\0\\\frac{1}{3}\end{bmatrix}$

$\nabla f = \begin{bmatrix}1\\-1\\1\end{bmatrix} \quad \nabla f^t d = -\frac{2}{3} < 0$

$Z(a) \neq \emptyset$

$a = (0, 1, 1/3) \quad d = \begin{bmatrix}d_1 \geq 0\\d_2 \leq 0\\d_3\end{bmatrix}$ where $\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}d_1\\d_2\\d_3\end{bmatrix} = 0$

$d_3 = \frac{-d_1 - 2d_2}{3}$

$D(a) = \mathcal{D}(a) = \begin{bmatrix}d_1 \geq 0\\d_2 \leq 0\\\frac{-d_1-2d_2}{3}\end{bmatrix} = \begin{bmatrix}d_1 \geq 0\\d_2 \leq 0\\d_3\end{bmatrix}$

$\text{Cone}\left(\begin{bmatrix}1\\0\\-\frac{1}{3}\end{bmatrix}, \begin{bmatrix}0\\-1\\\frac{2}{3}\end{bmatrix}\right)$

a) Consider any point a on the plate which is not on the line segment joining $(0, 0, 1)$ and $(0, 1, 1/3)$.

b) Then you have $a(1) > 0$ and you can move on the plate a little bit from a decreasing x-coordinate while keeping y-coordinate constant. What is the direction of your movement?

$\begin{bmatrix}1 & -1 & 1\end{bmatrix}\underset{v_1}{\begin{bmatrix}1\\0\\-\frac{1}{3}\end{bmatrix}} = 2/3 \quad \begin{bmatrix}1 & -1 & 1\end{bmatrix}\underset{v_2}{\begin{bmatrix}0\\-1\\2/3\end{bmatrix}} = 5/3 > 0$ | so for any non neg comb d of v_1 & v_2 we have $\nabla f^t d \geq 0$. so $Z(a) = \emptyset$

Answer. See, we are choosing $d = \begin{bmatrix}-1\\0\\z\end{bmatrix}$, where d must satisfy $\triangledown h^t(a)d = 0$ that is, $\begin{bmatrix}1 & 2 & 3\end{bmatrix}d = 0$. So $z = \frac{1}{3}$.

c) Is $\triangledown f(a)^t d < 0$ for that direction?

Answer. Yes, $D_d f(a) = -\frac{2}{3}$.

d) Is $Z(a) = \emptyset$? Can such points a be KT points?

Answer. No. No.

e) Next, assume that $a(1) = 0$ and the second coordinate $a(2) < 1$. (So, you are somewhere on the left most side except the point $(0, 1, 1/3)$.)

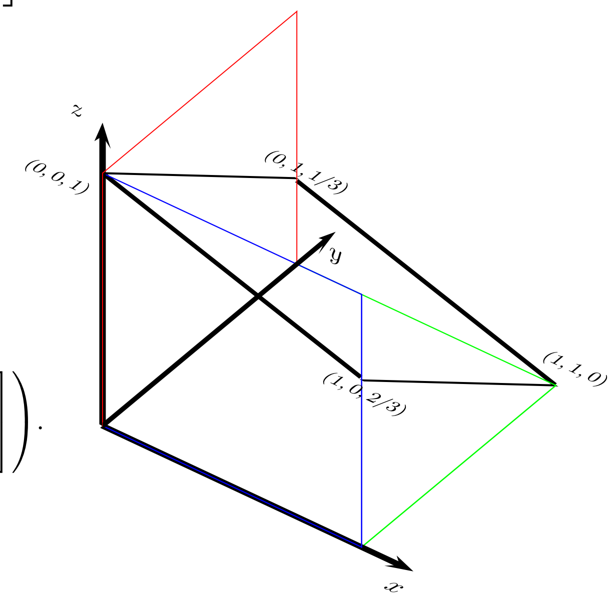
f) Can you give a feasible (?) direction d, along which $D_d f(a) < 0$?

Answer.

First of all, we are supposed to talk about the linearizing cone directions. But just to reject a, any feasible direction will do the job. And in any case, this set T is defined with all linear constraints. So at any point a, we have $D(a) = \mathcal{D}(a)$.

Coming back to the question, take the direction $d = \begin{bmatrix} 0 \\ 1 \\ -\frac{2}{3} \end{bmatrix}$ of movement along that line towards $(0, 1, 1/3)$. Then $D_d f(a) = -\frac{5}{3} < 0$.

g) Can such points a be KT points?



Answer. No.

h) So the only point that remains is $a = (0, 1, 1/3)$.

i) We already know that $\mathcal{D}(a) = \overline{D}(a) = D(a)$.

j) Argue that $\quad D(a) = \mathsf{cone}\left(d_1 = \begin{bmatrix} 1 \\ 0 \\ -1/3 \end{bmatrix}, d_2 = \begin{bmatrix} 0 \\ -1 \\ 2/3 \end{bmatrix} \right).$

Answer. Any direction must have the form $d = \begin{bmatrix} d_1 \geq 0 \\ d_2 \leq 0 \\ d_3 \end{bmatrix}$ and it must lie on the plane $x + 2y + 3z = 0$. Hence $d_3 = \frac{-d_1 - 2d_2}{3}$. Hence

$$d = \begin{bmatrix} d_1 \geq 0 \\ d_2 \leq 0 \\ \frac{-d_1 - 2d_2}{3} \end{bmatrix} = d_1 \begin{bmatrix} 1 \\ 0 \\ -\frac{1}{3} \end{bmatrix} - d_2 \begin{bmatrix} 0 \\ 1 \\ -\frac{2}{3} \end{bmatrix}$$

k) Do we have $D_{d_1} f(a) > 0$ and $D_{d_2} f(a) > 0$?

Answer. Yes, $D_{d_1} f(a) = 2/3 > 0$ and $D_{d_2} f(a) = 5/3 > 0$.

l) Can we conclude that $D_d f(a) \geq 0$ for each $d \in \mathcal{D}(a)$?

Answer. Yes. As $D(a) = \mathcal{D}(a)$.

m) So $Z(a) = \emptyset$ and $a = (0, 1, 1/3)$ is the only KT point.

`(Second (algebraic) way.)` Recall the problem

$$\begin{array}{ll} \min & x - y + z \\ \hline \text{s.t.} & g_1 \equiv 1 - x \geq 0, g_2 \equiv 1 - y \geq 0, g_3 \equiv 1 - z \geq 0, g_4 \equiv x \geq 0, \\ & g_5 \equiv y \geq 0, g_6 \equiv z \geq 0, h_1 \equiv x + 2y + 3z - 3 = 0 \end{array}$$

NEW

a) Suppose that ~~a feasible point~~ a is a KT point. So there exist $\lambda_i \geq 0, w \in \mathbb{R}$ with $\lambda_i g_i(a) = 0$ such that

$\lambda_2 = 0$
$\Rightarrow \omega \leq -\frac{1}{2}$
$\Rightarrow \lambda_4 > 0, \lambda_6 > 0$
$\Rightarrow a = (0, *, 0)$
$= (0, \frac{3}{2}, 0)$
$\notin T$

~~argument~~

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \lambda_1 \cancel{\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}} + \lambda_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_3 \cancel{\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}} + \lambda_4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_5 \cancel{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} + \lambda_6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} + \lambda_4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_5 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \lambda_6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

b) Must $\lambda_2 > 0$? What do we get from here? yes. $\Rightarrow g_2(a) = 0 \Rightarrow a = (*, 1, *)$.
As $\lambda_5 g_5(a) = 0$, we get $\lambda_5 = 0$

Answer. Yes. As $\triangledown f(a)$ has second coordinate -1, we must have $\boxed{\lambda_2 > 0}$. As $\lambda_2 g_2(a) = 0$, we get that $\boxed{a(2) = 1}$. Also $\boxed{\lambda_5 = 0}$.

c) Can $\lambda_3 > 0$? What do get from here? $\lambda_3 > 0$ (using $\lambda_3 g_3(a) = 0$) $\Rightarrow a = (*, 1, 1) \notin T$
as ~~a is not on the~~ $h_1(a) \neq 0$.

Answer. Suppose that $\lambda_3 > 0$. As $\lambda_3 g_3(a) = 0$, we get $a(3) = 1$. Hence $a = (\geq 0, 1, 1) \notin T$. We get $\boxed{\lambda_3 = 0}$.

d) Can $\lambda_1 > 0$? $\lambda_1 > 0$ (use $\lambda_1 g_1(a) = 0$) $\Rightarrow x = 1$, $\lambda_4 = 0$

Answer. Suppose that $\lambda_1 > 0$. So $a(1) = 1$ and $\lambda_4 = 0$. So our equation is

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$w = 1 + \lambda_1 \geq 1$

Looking at the first coordinates, that $w = 1 + \lambda_1$. In that case, the third coordinate $3w + 3\lambda_1 > 1$. So this case is not possible. So $\boxed{\lambda_1 = 0}$.

e) So our equation is

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \lambda_2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

f) Can we have $w > \frac{1}{3}$? No.

$w \le \frac{1}{3} \Rightarrow \lambda_4 \geq 2/3$

$\lambda_4 g_4(a) \Rightarrow x = 0$

Answer. No. Looking at the third coordinates.

$a = (0, 1, \frac{1}{3}) \to$ KT pt

g) Is $\lambda_4 > 0$? What do we get from here?

$\lambda_1 = 0, \lambda_5 = 0, \lambda_3 = 0, \quad w = 1/3, \lambda_4 = \frac{2}{3}, \lambda_6 = 0, \lambda_2 = \frac{5}{3}$

Answer. Yes, looking at the first coordinates. We get that $a(1) = 0$. As the plane equation must be satisfied, we get $a = (0, 1, \frac{1}{3})$.

h) Is a a KT point?

Answer. Yes. With $\lambda = (0, 5/3, 0, 2/3, 0, 0)$ and $w = 1/3$, it satisfies the definition of a KT point.

Some exercises

[32.4] **Practice** Find all KT points for

$$\begin{array}{ll} \min & f(x) = x_1 + x_2 \\ \text{s.t.} & g_1(x) = x_1^3 - x_2 \geq 0, \ g_2(x) = x_1 \geq 0, \ g_3(x) = x_2 \geq 0. \end{array}$$

[32.5] **NoPen** Consider maximizing f subject to $g_i(x) \geq 0$, $i = 1, \ldots, m$ and $h_j(x) = 0$, $j = 1, \ldots, p$. Let a be a point of local maximum. Then under some regularity conditions $Z(a)$ must be empty. Which one is an appropriate expression for $Z(a)$?

a) $\{d \mid \langle \nabla f(a), d \rangle > 0, \ \langle \nabla g_i(a), d \rangle \geq 0 \text{ for } i \in A(a), \ \langle \nabla h_j(a), d \rangle = 0 \text{ for each } j\}$

b) $\{d \mid \langle \nabla f(a), d \rangle < 0, \ \langle \nabla g_i(a), d \rangle \geq 0 \text{ for } i \in A(a), \ \langle \nabla h_j(a), d \rangle = 0 \text{ for each } j\}$

c) $\{d \mid \langle \nabla f(a), d \rangle > 0, \ \langle \nabla g_i(a), d \rangle \geq 0 \text{ for all } i, \ \langle \nabla h_j(a), d \rangle = 0 \text{ for each } j\}$

d) $\{d \mid \langle \nabla f(a), d \rangle < 0, \ \langle \nabla g_i(a), d \rangle \geq 0 \text{ for all } i, \ \langle \nabla h_j(a), d \rangle = 0 \text{ for each } j\}$

[32.6] **NoPen** Consider maximizing f subject to $g_i(x) \geq 0$, $i = 1, \ldots, m$ and $h_j(x) = 0$, $j = 1, \ldots, p$. Let a be a point of local maximum. Under some regularity conditions, a must be a KT point. Then which one of the following must be satisfied?

a) There exist $\lambda_i \geq 0$ and $w_j \in \mathbb{R}$, such that

$$\nabla f(a) = \sum_i \lambda_i \nabla g_i(a) + \sum_j w_j \nabla h_j(a) = 0, \qquad \lambda_i g_i(a) = 0, \forall i.$$

b) There exist $\lambda_i \geq 0$ and $w_j \in \mathbb{R}$, such that

$$-\bigtriangledown f(a) = \sum_i \lambda_i \bigtriangledown g_i(a) + \sum_j w_j \bigtriangledown h_j(a) = 0, \qquad \lambda_i g_i(a) = 0, \forall i.$$

[32.7] **Exercise(E)** Consider a set $T = \{x \mid g_1(x), \ldots, g_m(x) \geq 0, h_1, \ldots, h_p(x) = 0\}$. We want minimize a constant function on T. My friend thinks that each feasible point is a KT point. Is that true?