35 Lecture 35

KT necessary condition and sufficient conditions

The following is the result we were expecting earlier. This allows us to directly search for kt points in order to find local minimums, when ktcq1 holds on T.

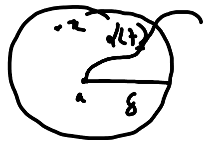
[35.1] **Lemma** (Kuhn-Tucker necessary condition (ktnc) for a local minimum) Let a be a local minimum of (P2) satisfying ktcq1. Then $Z(a) = \emptyset$ must hold, that is, a must be a kt point.

Proof. $d \in \mathcal{D}(a) \Rightarrow \nabla f(a)^t d \geq 0$. Let $d \in \mathcal{D}(a)$.

ktcq1 $\Rightarrow \exists\, \varepsilon > 0$ s.t. $\alpha : [0, \varepsilon] \to T$ s.t. $\alpha(0) = a$, $\lim\limits_{t \to 0+} \dfrac{\alpha(t) - a}{t} = d$

loc min $\Rightarrow \exists\, \delta > 0$, ~~$\delta < \varepsilon$~~ s.t. $x \in B_\delta(a) \cap T \Rightarrow f(x) \geq f(a)$

~~Let $x \in B_\delta(a) \cap T$~~ Let $\alpha(t) \in B_\delta(a)$.



$$0 \leq \frac{f(\alpha(t)) - f(\alpha(0))}{t} = \frac{f'(a + \theta(\alpha(t) - \alpha(0)))(\alpha(t) - \alpha(0))}{t}, \quad 0 < \theta < 1$$

$$0 \leq \lim_{t \to 0+} f'(a + \theta(\alpha(t) - \alpha(0))) \frac{\alpha(t) - \alpha(0)}{t} = f'(a)d$$

a) Let $d \in \mathcal{D}(a)$. As ktcq1 holds at a, there is a curve $\alpha(t)$, $t \in [0, \epsilon]$ such that $\alpha(0) = a$ and $\alpha(t) \in T$, $\forall t \in [0, \epsilon]$ and $d = \lim\limits_{t \to 0+} \frac{\alpha(t) - a}{t}$.

b) As a is a local minimum, $\exists \delta > 0$, $\delta \cancel{<} \epsilon$ such that $f(a) \leq f(x)$, $\forall x \in T \cap B_\delta(a)$.

c) By Taylor's theorem, for each $t \in (0, \delta)$, we have (take t small s.t. $\alpha(t) \in B_\delta(a)$)

$$0 \leq f(\alpha(t)) - f(\alpha(0)) = f'\Big(\alpha(0)) + \theta(\alpha(t) - \alpha(0))\Big)(\alpha(t) - \alpha(0)),$$

for some $0 < \theta < 1$. Dividing by t and letting $t \to 0+$, we get

$$f'(a)d = \lim_{t \to 0+} f'\Big(\alpha(0)) + \theta(\alpha(t) - \alpha(0))\Big) \frac{\alpha(t) - \alpha(0)}{t} \geq 0.$$

Thus $Z(a) = \emptyset$. ∎

$\nabla L(a) = 0$ $\quad H_L(a)$ psd in $B_\delta(a)$

For h small

$\Rightarrow a$ is lc min for L.

$0 \leq L(a+h) - L(a) = f(a+h) - \sum \lambda_i g_i(a+h) - \cancel{\sum \omega_j h_j(a+h)}$

$- f(a) - \sum \lambda_i \cancel{g_i}(a)$

$f(a+h) - f(a) = \sum \lambda_i g_i(a+h) \geq 0 \Rightarrow f(a+h) \geq f(a)$
$\Rightarrow a$ is the min for f.

[35.2] **Theorem** (`Sufficient condition for a kt point to be a local minimum (ktsc)`) Let a be a kt point for (P2) with functions in $\mathcal{C}^2$. Let

i) $A_+(a) = \{i \in A(a) \mid \lambda_i > 0\}$, and $\to$ (super active constraints)
ii) $\mathcal{D}_+(a) = \{d \in \mathcal{D}(a) \mid \nabla g_i(a)^t d = 0$ for all $i \in A_+(a)\}$. (Thus $\mathcal{D}_+(a) = \mathcal{D}(a)$ if $A_+(a) = \emptyset$.)

Recall that $L \equiv f - \sum_i \lambda_i g_i - \sum_j w_j h_j$ is the Lagrangian function. Then the following are true.

✓ a) If $d^t H(L)_a d > 0$ for each nonzero $d \in \mathcal{D}_+(a)$, then a is a strict local minimum.
✓ b) If $H(L)_a$ is pd, then a is strict local minimum.
✓ c) If $H(L)$ is psd in some $B_\delta(a)$, then a is a local minimum.

Proof. a) suppose $d^t H_L d > 0 \ \forall d \in \mathcal{D}_+(a), d \neq 0$.

suppose it is not. $\exists a_k \to a$ s.t. $a_k \neq a$, $f(a_k) \leq f(a)$.

$a_k - a = \|a_k - a\| \dfrac{a_k - a}{\|a_k - a\|} = t_k d_k$ | $\|d_k\| = 1$, $(d_1, d_2, \dots)$, wlog let $d_k \to d$.

(1) $0 \geq f(a_k) - f(a) = t_k \nabla f(a)^t d_k + t_k^2 d_k^t H_f(a) d_k + o(t_k^2)$

g_i active $\to$ (2) $0 \leq g_i(a_k) - g_i(a) = t_k \nabla g_i(a)^t d_k + t_k^2 d_k^t H_{g_i}(a) d_k + o(t_k^2)$

(3) $0 = h_j(a_k) - h_j(a) = t_k \nabla h_j(a)^t d_k + t_k^2 d_k^t H_{h_j}(a) d_k + o(t_k^2)$

$$0 \geq t_k \Big[\nabla f(a)^t - \sum_{i \in A(a)} \lambda_i \nabla g_i(a)^t - \sum w_j \nabla h_j(a)^t\Big] d_k$$
$$+ t_k^2 d_k^t H_L d_k + o(t_k^2)$$
$$= t_k \underbrace{\nabla L(a)^t}_{=0} d_k + t_k^2 d_k^t H_L(a) d_k + o(t_k^2)$$
$$= t_k^2 d_k^t H_L(a) d_k + o(t_k^2)$$

$$\frac{0}{t_k^2} \geq d_k^t H_L(a) d_k + \frac{o(t_k^2)}{t_k^2} \to 0, \quad \text{take } k \to \infty$$

we get $0 \geq d^t H_L(a) d$.

From (2) & (3), we get $d \in \mathcal{D}(a)$.

As $d^t H_L(a) d \leq 0$, we see that $d \notin \mathcal{D}_+(a)$. So $\exists$ a super-active g_i s.t. $\nabla g_i(a)^t d > 0$

$$\nabla f(a)^t d = \underbrace{\sum \lambda_i \nabla g_i(a)^t d}_{> 0} + \sum \omega_j \underbrace{\nabla h_j(a)^t d}_{0}$$

Look at ①, divide by t_k, take $k \to \infty$, we get $\nabla f(a)^t d \leq 0$. $\Rightarrow \Leftarrow$

a) Suppose that a is not a local minimum. Then $\exists$ a sequence $a_k \to a$, $a_k \neq a$ and $f(a_k) \leq f(a)$. Write

$$a_k = a + (a_k - a) = a + \|a_k - a\| \frac{a_k - a}{\|a_k - a\|} = a + t_k d_k.$$

As $a_k \to a$, we have $t_k = \|a_k - a\| \to 0$. As $(d_k = \frac{a_k - a}{\|a_k - a\|})$ is a bounded sequence (being unit vectors), and the unit sphere is compact, it has a convergent subsequence converging to (a unit vector) d, say. Without loss, we assume that d_k itself is converging to d. So, using Taylor's theorem, we get

$$\begin{array}{rclcl} 0 & \geq & f(a_k) - f(a) & = & t_k \bigtriangledown f(a)^t d_k + \frac{1}{2} t_k^2 d_k^t H_f d_k + o(t_k^2) \\ 0 & = & h_j(a_k) & = & t_k \bigtriangledown h_j(a)^t d_k + \frac{1}{2} t_k^2 d_k^t H_{h_j} d_k + o(t_k^2) \\ 0 & \leq & g_i(a_k) & = & t_k \bigtriangledown g_i(a)^t d_k + \frac{1}{2} t_k^2 d_k^t H_{g_i} d_k + o(t_k^2),\ i \in A(a). \end{array}$$

We get (as a is a KT point)

$$0 \geq t_k \bigtriangledown L^t d_k + \frac{1}{2} t_k^2 d_k^t H_L d_k + o(t_k^2) = \frac{1}{2} t_k^2 d_k^t H_L d_k + o(t_k^2).$$

Dividing by t_k^2 and taking limit, we see that $d^t H_l d \leq 0$. Hence $d \notin \mathcal{D}_+(a)$. Dividing last two by t_k and taking the limit, we see that $\bigtriangledown h_j^t d = 0$ and $\bigtriangledown g_i^t d \geq 0$ for $i \in A(a)$ and so $d \in \mathcal{D}(a)$. That is, $\exists i \in A_+(a)$ such that $\bigtriangledown g_i^t d > 0$. But, then $\bigtriangledown f^t d > 0$ as $\bigtriangledown L = 0$. But, now dividing the first inequation by t_k and taking limit, we see that $\bigtriangledown f^t d \leq 0$. This is a contradiction.

The proof of b) follows from a).

To prove c) note that as $\bigtriangledown L = 0$, and H_L is psd in a neighborhood, we see that a is a point of local minimum of L. Thus

$$0 \leq L(a+h) - L(a) = f(a+h) - \sum \lambda_i g_i(a+h) - f(a) - \sum \lambda_i g_i(a) = f(a+h) - \sum \lambda_i g_i(a+h) - f(a),$$

as $\lambda_i g_i(a) = 0$ for each i. So

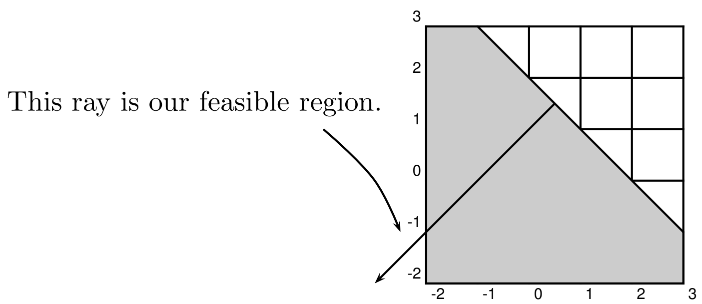
$$f(a+h) - f(a) \geq \sum \lambda_i g_i(a+h) \geq 0.$$

Thus a is a point of local minimum for f. ■

[35.3] **Exercise(E)** Write an alternate proof of part b) of [35.2] in the line of the proof of part c).

[35.4] **Example** Consider

$$\min \ f(x) = (x_1-1)^2 + x_2 - 2$$
$$\text{s.t.} \ g(x) = 2 - x_1 - x_2 \geq 0, \ h(x) = x_2 - x_1 - 1 = 0.$$



∘ Constraints are linear. So ktcq1 holds at each feasible point. So by ktnc, each point of local minimum is a kt point.

∘ We find kt points: $\begin{bmatrix} 2(x_1-1) \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ -1 \end{bmatrix} + w \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\lambda(2 - x_1 - x_2) = 0$.

∘ If $\lambda = 0$, then $w = 1$. Equating the first row we get $x_1 = .5$. Applying $h(x)$, we get $x_2 = 1.5$. Thus $(.5, 1.5)$ being feasible, is a KT point with $\lambda = 0, w = 1$.

∘ If $\lambda > 0$, then $x_1 + x_2 = 2$. Using $h(x)$, we get $x_1 = .5$. But as $w > 1$, we must have $2x_1 < 1$, not possible. So we do not get any kt point here.

∘ Going with $\lambda = 0$, $w = 1$, we have $A_+(a) = \emptyset$. So $\mathcal{D}_+(a) = \mathcal{D}(a) = \{d \mid d_1 + d_2 \leq 0, d_1 = d_2\}$. We have $H(L)_a = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$. For any nonzero $d = \begin{bmatrix} -t \\ -t \end{bmatrix}$, with $t > 0$, we have $d^t H(L)_a d = 2t^2 > 0$. By ktsc, a is a strict local minimum.

[35.5] **Example** `(One dimensional example.)` Consider $\begin{array}{ll} \min & \underline{x^2} \\ \text{s.t.} & x \ge 5, x \le 6 \end{array}$ and $\begin{array}{ll} \max & \underline{x^2} \\ \text{s.t.} & x \ge 5, x \le 6 \end{array}$.

Find the kt points for both the problems apply ktsc.

Answer.

a) First consider the minimization problem. We have $L = x^2 - \lambda_1(x-5) - \lambda_2(6-x)$.

b) Constraints are linear. So ktcq1 holds at all feasible points. Any local minimum must be a kt point.

c) Find kt points:

$$2x = \lambda_1 - \lambda_2.$$

If $\lambda_2 > 0$, then $x = 6$. As $\lambda_1 g_1(a) = 0$, we must have $\lambda_1 = 0$. But then $12 = 0 - \lambda_2$ is not possible.

So $\lambda_2 = 0$. If $\lambda_1 > 0$, then $x = 5$. Notice that $x = 5$ is a kt point with $\lambda_1 = 10$ and $\lambda_2 = 0$. Notice that $H(L) = [2]$ is pd. Hence the point $x = 5$ is a strict local minimum by ktsc.

A) Consider the maximization problem, that is, minimizing $-x^2$. Approaching similarly, we get $x = 6$ as a kt point.

B) However, $H(L) = -2$. So we cannot use part b) of ktsc.

C) However, we can use part a). Note that $A_+(6) = \{2\}$ and $\mathcal{D}_+(6) = \{0\}$. Hence, for each nonzero $d \in \mathcal{D}_+(6)$ we have $d^t H(L) d > 0$. So $x = 6$ satisfies ktsc part a), to be a point of strict minimum for $-x^2$.

Exercises

[35.6] **NoPen** Let a be any feasible point for $\begin{array}{ll} \max & \underline{f(x)} \\ \text{s.t.} & \begin{array}{l} g_i(x) \ge 0,\ i = 1, \ldots, m, \\ h_j(x) = 0,\ j = 1, \ldots, p. \end{array} \end{array}$

Suppose that $\mathcal{D}(a) \cap \{d \mid \bigtriangledown f(a)^t d > 0\} = \emptyset$. Is it necessary that a is a KT point?

[35.7] **Exercise(M)** Let a be a local minimum for

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \geq 0,\ i = 1, \ldots, m,\ \ h_j(x) = 0,\ j = 1, \ldots, p,\ \ x_k \geq 0, k = 1, \ldots, n. \end{array}$$

a) Assume that $Z(a) = \emptyset$. Let

$$L(x, \lambda, w) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) - \sum_{j=1}^{p} w_j h_j(x).$$

Prove that $\exists\ \lambda_i \geq 0,\ i = 1, \ldots, m,\ w_j,\ j = 1, \ldots, p$ such that the following KT conditions are satisfied

$$\bigtriangledown L(a, \lambda, w) \geq 0, \qquad \lambda_i g_i(a) = 0, \forall i = 1, \ldots, m, \qquad a^t \bigtriangledown L(a, \lambda, w) = 0.$$

b) Is the converse of a) true?

c) Is $a = (0, \sqrt{2}, \sqrt{2})$ a KT point of

$$\begin{array}{ll} \min & f = x_1^3 - 6x_1^2 + 11x_1 + x_3 \\ \text{s.t.} & g_1 \equiv -x_1^2 - x_2^2 + x_3^2 \geq 0,\ \ g_2 \equiv x_1^2 + x_2^2 + x_3^2 - 4 \geq 0,\ \ g_3 \equiv 5 - x_3 \geq 0,\ x_i \geq 0 \end{array} ?$$

[35.8] **Practice** Consider the problem $\begin{array}{ll} \text{opt} & x_1 + x_2^2 + x_3^3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1,\ x_i \geq 0. \end{array}$

a) Consider the minimization problem first and let $a = (0, 1, 0)$. Compute $D(a)$, $\mathcal{D}(a)$, $Z(a)$. Does ktcq1 hold at a? Is a a kt point? Is it a local minimum?

b) Can $(0, 1, 0)$ be a point of local maximum?

c) Does ktcq1 hold for every feasible point?

[35.9] **Practice** Consider the problem $\begin{array}{ll} \max & x^2 + y \\ \text{s.t.} & x \geq 0,\ \ x^2 + y^2 = 1. \end{array}$

a) Do you think, at each $a \in T$, KTCQ1 holds?

b) Find the KT points.

c) Conclude, whether these points are local maxima.

[35.10] **Practice** Find all local minimums for $\begin{array}{ll} \min & x^2 + y^2 + z^2 - x - y - z \\ \text{s.t.} & 0 \leq x, y, z \leq 1. \end{array}$

[35.11] **Exercise(E)** Consider the problems $\begin{array}{ll} \min & x^4 \\ \text{s.t.} & |x| \leq 1 \end{array}$ and $\begin{array}{ll} \min & x^3 \\ \text{s.t.} & |x| \leq 1 \end{array}$.

a) Argue that $a = 0$ is a KT point for both.

b) Now show that ktsc part a) is not a necessary condition. Hence, conclude that *if* $d^t H(L)(a) d = 0$ *for some nonzero* $d \in \mathcal{D}_+(a)$, *then* a *may or may not be a point of local minimum.*

c) However, show that we can still apply the ktsc part c) to conclude that 0 is a point of local minimum for the first One.

[35.12] **Exercise(M)** Consider the problem $\begin{array}{ll} \max & \underline{x_1 x_2 x_3 x_4} \\ \text{s.t.} & x_1 + x_2 + x_3 + x_4 = 5,\ x_i \geq 0. \end{array}$ We already know one way to find the absolute maximum. Find all KT points and give another solution using KT theory. Also search for all local maximums.

[35.13] **Practice** Consider minimizing $x^2 + y^2 + xy$ over the region $T = \{x^2 + y^2 \geq 1,\ x^2 + 4y^2 \leq 4\}$. Show that a point of local minimum must be a KT point.

[35.14] **Practice** Consider minimizing $x^2 + y^2 + xy$ over the region $T = \{x^2 + y^2 \geq 1,\ x^2 + y^2 \leq 4\}$. Assume that the points of local minimums are KT points. Find all KT points. Apply KTSC at these points.