The KKT conditions

Lecture 16

November 30, 2021

Beck, sections 11.1-11.3

Inequality constrained optimization

Goal: solve the optimization problem:

$$\min f(\underline{x})$$
$$\text{s.t. } g_i(\underline{x}) \leq 0 \quad \forall\ i = 1 \ldots m$$

where $\underline{x} \in \mathbb{R}^n$, $f, g_i \in C^1$.

Note: Inequality constraints are (much) more "difficult" to handle than equality constraints.

Ex: For equality-constrained problems, we already have a tool to solve these: Lagrange multipliers.

KKT conditions: (Karush-Kuhn-Tucker)

"Necessary conditions for local optimality"

With the above notation, let $\underline{x}^*$ be a local minimum. Assume that $\{\nabla g_i(\underline{x}^*)\}_{i=1}^{m}$ is a collection of linearly independent vectors*.

Then:

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\underline{x}^*) = \underline{0}$$ for some $\underline{\lambda} = (\lambda_1, \ldots \lambda_m) \in \mathbb{R}_+^m$ that is nonzero.

also: $\lambda_i g_i(\underline{x}^*) = 0 \quad \forall\, i = 1, \ldots m.$

*: This condition is actually more strict than necessary.

Necessary optimality conditions

Definition: Consider the optimization problem

$$\min f(\underline{x}) \text{ s.t. } \underline{x} \in C,$$

where $f \in C^1$ and C is closed and convex. A vector $\underline{d} \in \mathbb{R}^n$ is a feasible descent direction at $\underline{x}$ if $\nabla f(\underline{x})^T \underline{d} < 0$ and if $\exists\, \varepsilon > 0$ s.t. $\underline{x} + t\underline{d} \in C \;\; \forall\, t \in [0, \varepsilon]$

Lemma: Consider the optimization problem

$$\min f(\underline{x}) \text{ s.t. } \underline{x} \in C.$$

A necessary condition for $\underline{x}^*$ to be a local minimum is that there are no feasible descent directions at $\underline{x}^*$.

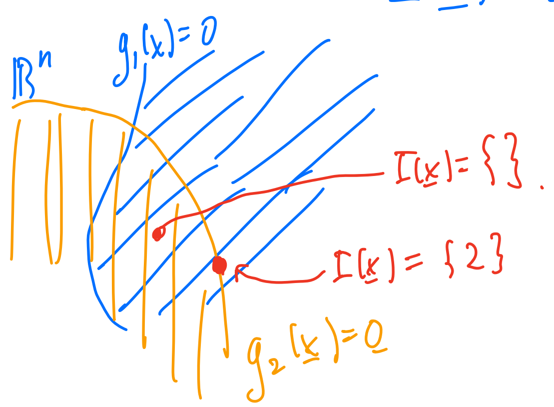
<u>Proof</u> (sketch) By contradiction: if $\underline{x}^*$ is a local min and $\underline{d}$ is a feasible descent direction, then we can take a small step in direction $\underline{d}$ and decrease value of f, violating definition of a local minimum.

Back to inequality-constrained problem:

$$\min f(\underline{x}) \quad \text{s.t.} \quad g_i(\underline{x}) \le 0 \;, \; i=1..m$$

We will need a notion "active constraints". At a point $\underline{x} \in \mathbb{R}^n$, the set of active constraint indices is

$$I(\underline{x}) = \{ i \in \{1,2,..,m\} \text{ s.t. } g_i(\underline{x}) = 0 \}$$



<u>Lemma</u>: Consider the optimization problem

$$\min f(\underline{x}) \quad \text{s.t.} \quad g_i(\underline{x}) \le 0 \;, \; i=1..m.$$

Assume $f, g_i \in C^1$. Let $\underline{x}^*$ be a local minimum.

Then there does <u>not</u> exist a vector $\underline{d} \in \mathbb{R}^n$ s.t.

$$\nabla f(\underline{x}^*)^T \underline{d} < 0$$
$$\nabla g_i(\underline{x}^*)^T \underline{d} < 0 \quad \forall \ i \in I(\underline{x}^*).$$

<u>Proof</u>: (sketch) Again by contradiction. Assume there does exist a $\underline{d}$ s.t. $\nabla f(\underline{x}^*)^T \underline{d} < 0$ and $\nabla g_i(\underline{x}^*)^T \underline{d} < 0 \ \forall \ i \in \underline{I(x^*)}$.

We can choose an $\varepsilon > 0$ s.t. $\underline{x}^* + \varepsilon \underline{d}$ accomplishes:

(☆☆) (i) $f(\underline{x}^* + \varepsilon \underline{d}) < f(\underline{x}^*)$

(☆) $\begin{cases} \text{(ii) } g_i(\underline{x}^* + \varepsilon \underline{d}) < g_i(\underline{x}^*) = 0 \quad \forall \ i \in I(\underline{x}^*) \\ \text{(iii) } \forall \ i \notin I(\underline{x}^*) \text{ can choose } \varepsilon \text{ small enough so that } g_i(\underline{x}^* + \varepsilon \underline{d}) < 0 \text{ since } g_i(\underline{x}^*) < 0. \end{cases}$

(☆): $\underline{x}^* + \varepsilon \underline{d}$ is feasible.

(☆☆): $\underline{x}^* + \varepsilon \underline{d}$ has smaller value of f than $\underline{x}^*$

$\longrightarrow$ contradicts def'n of $\underline{x}^*$ being a local min.

Farkas' Lemma and Gordan's Theorem

These two results are examples of "theorems of the alternative".

Farkas' Lemma: Let $\underline{\underline{A}} \in \mathbb{R}^{m \times n}$, $\underline{b} \in \mathbb{R}^m$. Then exactly one of the following is true:

non-neg. orthant in $\mathbb{R}^n$

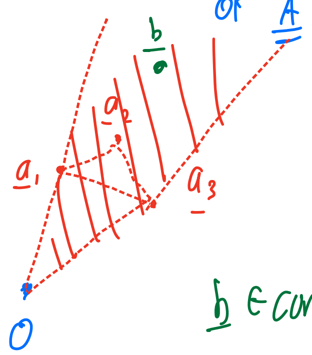
(i) $\exists\ \underline{x} \in \mathbb{R}^n_+$ s.t. $\underline{\underline{A}}\,\underline{x} = \underline{b}$ $(\underline{x} \neq \underline{0})$

(ii) $\exists\ \underline{y} \in \mathbb{R}^m$ s.t. $\underline{y}^T \underline{\underline{A}} \geq \underline{0}$ and $\underline{y}^T \underline{b} < 0$

component-wise.

Interpretation: (i) $\underline{b}$ is a conic combination of the columns of $\underline{\underline{A}}$. ($\underline{b}$ is in the conic hull of the columns of $\underline{\underline{A}}$).

$\mathbb{R}^m$



$\underline{b} \in \text{cone}(\underline{a}_1 .. \underline{a}_n)$

$$\underline{\underline{A}} = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix}$$

(ii) $\exists\, y \in \mathbb{R}^m$ defining a hyperplane in $\mathbb{R}^m$

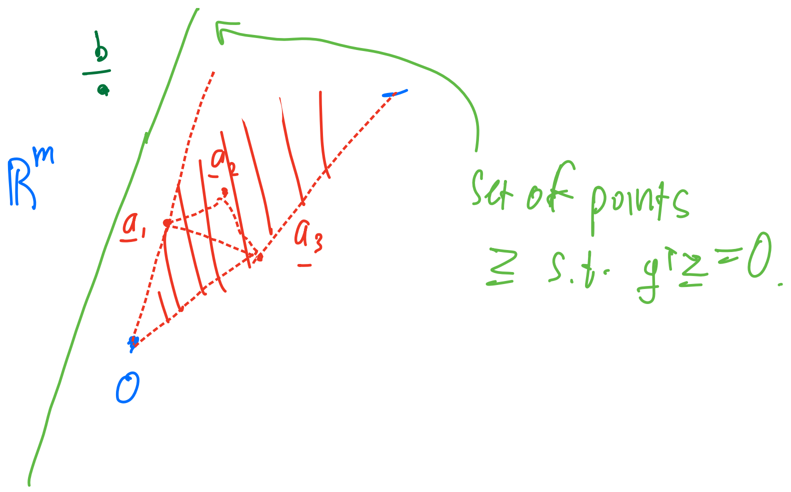
(set of $\underline{z}$ s.t. $\underline{z}^T y = 0$)

We have $y^T b < 0$: $\underline{b}$ lies on one side of the hyperplane

$$y^T \underline{\underline{A}} \geq 0$$

$$(y^T \underline{a}_1,\ y^T \underline{a}_2 \ \ldots\ y^T \underline{a}_n) \geq 0$$

We have $y^T \underline{\underline{A}} \geq 0$: all columns of $\underline{\underline{A}}$ lie on the other side of hyperplane.



Gordan's Theorem ("of the alternative")

Let $\underline{\underline{A}} \in \mathbb{R}^{m \times n}$. Then exactly one of the following is true:

(i) $\exists\, \underline{x} \in \mathbb{R}^n_+$, nonzero, s.t. $\underline{\underline{A}}\,\underline{x} = \underline{0}$

(ii) $\exists\, y \in \mathbb{R}^m$ s.t. $y^T \underline{\underline{A}} < 0$.

Proof is almost direct from Farkas' Lemma taking $\underline{b} = \underline{0}$.
(Farkas' lemma yields $y^T \underline{\underline{A}} \leq 0$, but can make this a strict inequality.)

Back to constrained optimization

The ~~Fritz-John~~ conditions

Fritz John

<u>Theorem</u>: Consider the problem:

$$\min f(\underline{x}) \quad \text{s.t.} \quad g_i(\underline{x}) \le 0 \quad \forall\, i = 1 .. m$$

$f, g_i \in C^1$. Let $\underline{x}^*$ be a local minimum. Then

$\exists\ \underline{\lambda} = (\lambda_0, \lambda_1, \ldots \lambda_m) \in \mathbb{R}^{m+1}_+$ s.t. $(\underline{\lambda} \neq \underline{0})$

$$\lambda_0 \nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\underline{x}^*) = 0.$$

also: $\lambda_i g_i(\underline{x}^*) = 0 \quad \forall\, i = 1 .. m.$

Proof: By a previous Lemma: $\nexists\ \underline{d} \in \mathbb{R}^n$ s.t.

$$\begin{cases} \nabla f(\underline{x}^*)^T \underline{d} < 0 \\ \nabla g_i(\underline{x}^*)^T \underline{d} < 0 \quad \forall\ i \in I(\underline{x}^*) \end{cases}$$

$$\underline{d}^T \overbrace{\begin{pmatrix} \nabla f(\underline{x}^*) & \nabla g_{j_1}(\underline{x}^*) & \ldots & \nabla g_{j_k}(\underline{x}^*) \end{pmatrix}}^{\text{"}A\text{"}} < \underline{0}$$

$\{j_1 \ldots j_k\} = I(\underline{x}^*)$

$\xRightarrow{\text{Gordan's Theorem}} \exists\ \underline{\lambda} \in \mathbb{R}_+^{m+1}$ s.t. $\underline{\underline{A}}\,\underline{\lambda} = \underline{0}$ $(\underline{\lambda} \neq \underline{0})$

$$\underline{\underline{A}}\,\underline{\lambda} = \lambda_0 \nabla f(\underline{x}^*) + \sum_{i \in I(\underline{x}^*)} \lambda_i \nabla g_i(\underline{x}^*) = \underline{0}$$

$\lambda_i g_i(\underline{x}^*) = 0$
$\forall\ i$

Now: Set $\lambda_i = 0$ when $i \notin I(\underline{x}^*)$. ∎

We have: $g_i(\underline{x}^*) = 0$ $i \in I(\underline{x}^*)$, $\lambda_i = 0$, $i \notin I(\underline{x}^*)$

Fritz John Conditions are not as useful as one might think: $\lambda_0 = 0$ can happen.

Next Thursday: "review" session.

Final exam: comprehensive exam
120 minutes
No notes/calculators.
Content based on HW. (1-6)
Wed Dec. 15 @ 8am (here)

Office hours: until Dec 10 (next Friday), office hours as usually scheduled.
Tues Dec 14, office hours from 11-12pm.

HW solutions: HW 1-5 solutions available (maybe 5 right after class.)
HW 6 solutions: available end of day on Dec 10 (Friday)

Please fill out course evals!

KKT conditions: inequality constrained problems

<u>Theorem</u> Consider

$$\min f(\underline{x}) \text{ s.t. } g_i(\underline{x}) \leq 0 \quad \forall\, i = 1, \ldots, m.$$

Assume $f, g_i \in C^1$ and $\underline{x}^*$ is a local minimum. Assume $\{\nabla g_i(\underline{x}^*)\}_{i \in I(\underline{x}^*)}$ are linearly independent. Then $\exists\, \underline{\lambda} \in \mathbb{R}^m_+$, $\underline{\lambda} \neq \underline{0}$ s.t.

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\underline{x}^*) = \underline{0}.$$

also: $\lambda_i g_i(\underline{x}^*) = 0, \quad i = 1, \ldots, m$

How we got from Fritz John $\rightarrow$ KKT?

Pf (KKT): By Fritz-John conditions $\exists\ \underset{\sim}{\eta} \in \mathbb{R}_+^{m+1}$ s.t.

$$\eta_0 \nabla f(\underset{\sim}{x}^*) + \sum_{i=1}^{m} \eta_i \nabla g_i(\underset{\sim}{x}^*) = 0, \quad \underset{\sim}{\eta} \neq \underline{0}$$

Assumed $\{\nabla g_i(\underset{\sim}{x}^*)\}_{i \in I(\underset{\sim}{x}^*)}$ are linearly independent.

Recall: $\forall\ i \notin I(\underset{\sim}{x}^*)$, $\eta_i = 0$.

$$\Longrightarrow \eta_0 \nabla f(\underset{\sim}{x}^*) + \sum_{i \in I(\underset{\sim}{x}^*)} \eta_i \nabla g_i(\underset{\sim}{x}^*) = \underline{0}, \quad \underset{\sim}{\eta} \neq \underline{0}.$$

Under the linear independence assumptions, $\eta_0 \neq 0$.

(If $\eta_0 = 0$, then $\sum_{i \in I(\underset{\sim}{x}^*)} \eta_i \nabla g_i(\underset{\sim}{x}^*) = \underline{0}$, $\underset{\sim}{\eta} \neq \underline{0}$).

$$\xrightarrow{1/\eta_0} \nabla f(\underset{\sim}{x}^*) + \sum_{i=1}^{m} \frac{\eta_i}{\eta_0} \nabla g_i(\underset{\sim}{x}^*) = 0$$

$\lambda_i \geq 0$, $\underline{\lambda} = (\lambda_1 \ldots \lambda_m) \neq \underline{0}$. ∎

KKT conditions: example

Example (Beck, 11.6(i))

Use the KKT conditions to compute an optimal solution to:

$$\min_{\boldsymbol{x}} 3x_1^2 + x_2^2 \quad \text{subject to} \quad x_1 - x_2 + 8 \leqslant 0, \;\; x_2 \geqslant 0.$$

$-x_2 \leq 0$

$f(\underline{x}) = 3x_1^2 + x_2^2$

$g_1(\underline{x}) = x_1 - x_2 + 8$

$g_2(\underline{x}) = -x_2$

$$\underbrace{\nabla g_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \nabla g_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}}_{}$$

linearly independent $\forall\, \underline{x}$

$\Longrightarrow$ KKT conditions are necessary.

Which points satisfy KKT conditions?

$$\nabla f = \begin{pmatrix} 6x_1 \\ 2x_2 \end{pmatrix} \quad \nabla g_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \nabla g_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\left.\begin{aligned} &\nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0 \\ &\lambda_1 g_1 = 0 \\ &\lambda_2 g_2 = 0 \\ &\lambda_1, \lambda_2 \geq 0 \end{aligned}\right\} \rightarrow \begin{aligned} &6x_1 + \lambda_1 = 0 \\ &2x_2 - \lambda_1 - \lambda_2 = 0 \\ &\lambda_1 (x_1 - x_2 + 8) = 0 \\ &\lambda_2 (-x_2) = 0 \end{aligned}$$

$\lambda_1 = \lambda_2 = 0$: can't happen ($\underline{\lambda} \neq \underline{0}$)

<u>$\lambda_1 \neq 0$, $\lambda_2 = 0$</u> : $x_1 - x_2 + 8 = 0$

$$\left.\begin{aligned} &6x_1 + \lambda_1 = 0 \\ &2x_2 - \lambda_1 = 0 \end{aligned}\right\rangle \; 6x_1 + 2x_2 = 0 \rightarrow x_2 = -3x_1$$

$x_1 - x_2 + 8 = 0$

$x_1 - (-3x_1) + 8 = 0$

$x_1 = -2$

$x_2 = 6$

$\lambda_1 = 12$

$(x_1, x_2, \lambda_1, \lambda_2) = (-2, 6, 12, 0)$.

$\underline{\lambda_1 = 0, \quad \lambda_2 \neq 0}$:

$$\left.\begin{array}{l} 6x_1 + \lambda_1 = 0 \\ 2x_2 - \lambda_1 - \lambda_2 = 0 \\ \lambda_1 (x_1 - x_2 + 8) = 0 \\ \lambda_2 (-x_2) = 0 \end{array}\right\} \quad \begin{array}{l} 6x_1 = 0 \\ 2x_2 - \lambda_2 = 0 \\ \\ x_2 = 0 \end{array}$$

$\downarrow$

no solution

$(\lambda_2 \neq 0)$

$\underline{\lambda_1 \neq 0, \lambda_2 \neq 0}$:

$$x_2 = 0$$

$$x_1 - x_2 + 8 = 0 \rightarrow x_1 = -8$$

$$\lambda_1 = -6x_1 = 48$$

$$\lambda_2 = 2x_2 - \lambda_1 = -48 \quad \times$$

(no solution)

Summary: only point satisfying KKT conditions is

$(x_1, x_2) = (-2, 6) \quad (\lambda_1 = 12, \lambda_2 = 0)$

I.e. $(x_1, x_2) = (-2, 6)$ is a local (also global) minimum.

The KKT conditions can be useless:

$$\min\ x \qquad f(x) = x$$

$$\text{s.t. } x^2 \le 0. \qquad g_1(x) = x^2 \qquad \nabla g = 2x.$$

At the feasible point $x=0$, $\nabla g = 0$

not linearly independent.

KKT conditions: in/equality constrained problems

<u>Theorem</u>: Consider the optimization problem

$$\min_{\underline{x}} f(\underline{x}) \quad \text{s.t.} \quad g_i(\underline{x}) \leq 0 \quad \forall\, i=1,\dots m$$
$$h_j(\underline{x}) = 0 \quad \forall\, j=1,\dots p,$$

where $f, g_i, h_j \in C^1$. Let $\underline{x}^*$ be a local minimum and assume $\{\nabla g_i(\underline{x}^*)\}_{i \in I(\underline{x}^*)} \cup \{\nabla h_j(\underline{x}^*)\}_{j=1}^{p}$ are linearly independent.

Then: $\exists$ $\underline{\lambda} \in \mathbb{R}^m_+$, $\underline{\lambda} \neq 0$, $\underline{\mu} \in \mathbb{R}^p$ s.t.

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\underline{x}^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(\underline{x}^*) = \underline{0}$$

$$\lambda_i g_i(\underline{x}^*) = 0 \quad \forall\, i = 1, \ldots m.$$

KKT points and regularity

Consider $\min_{\underline{x}} f(\underline{x})$. s.t. $g_i(\underline{x}) \le 0$, $i = 1 \ldots m$

$h_j(\underline{x}) = 0$, $j = 1, \ldots p$.

Definition: A feasible point $\underline{x}$ is a KKT point if

$\exists\ \underline{\lambda} \in \mathbb{R}^m_+$, $\underline{\lambda} \neq \underline{0}$, $\underline{\mu} \in \mathbb{R}^p$ s.t.

$$\nabla f(\underline{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\underline{x}) + \sum_{j=1}^{p} \mu_j \nabla h_j(\underline{x}) = 0$$

$$\lambda_i g_i(\underline{x}) = 0, \quad i = 1 \ldots m.$$

Definition: A feasible point $\underline{x}$ is regular if it satisfies: $\{\nabla g_i(\underline{x})\}_{i \in \mathcal{I}(\underline{x})} \cup \{\nabla h_j(\underline{x})\}_{j=1}^{p}$ are linearly independent.

KKT theorem: If a local min $\underline{x}^*$ is regular, then it's a KKT point.

Consider the KKT conditions: define

$$L(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(\underline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\underline{x}) + \sum_{j=1}^{p} \mu_j h_j(\underline{x}).$$

"Lagrangian"

$$\nabla f + \sum_{i=1}^{m} \lambda_i \nabla g_i + \sum_{j=1}^{p} \mu_j \nabla h_j = \underline{0} \longrightarrow \text{stationarity of Lagrangian}$$

$$\left.\begin{matrix} g_i(\underline{x}) \le 0 \\ h_j(\underline{x}) = 0 \end{matrix}\right\} \text{"primal feasibility"}$$

$\underline{\lambda} \in \mathbb{R}_+^m$: "dual feasibility"

$\lambda_i g_i(\underline{x}) = 0 \quad \forall\, i = 1 \ldots m$: "complementary slackness"

KKT conditions: convex optimization

KKT conditions are sufficient (!).

Theorem: Suppose f, g_i are convex and h_j are affine.

If $\underline{x}^*$ is a KKT point, then it's a global minimizer.

Pf: Let $\underline{x}$ be any other feasible point.

$\underline{x}^*$ is a KKT point: $h_j(\underline{x}^*) = 0$, $j = 1 \ldots p$

$\lambda_i g_i(\underline{x}^*) = 0$, $i = 1 \ldots m$

Lagrangian $L(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(\underline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\underline{x}) + \sum_{j=1}^{p} \mu_j h_j(\underline{x})$

convex $\lambda \geq 0$

0 for any feasible point.

$\Longrightarrow$ L is convex.

$\Longrightarrow$ stationary points are global minima.

KKT point: $\nabla L(\underline{x}^*, \lambda, \mu) = \underline{0}$.

$\Longrightarrow L(\underline{x}^*) \leq L(\underline{x})$

convexity of L and stationarity of KKT points.

$$f(\underline{x}^*) = L(\underline{x}^*) \leq L(\underline{x}) = f(\underline{x}) + \sum_{j=1}^{p} \mu_j h_j(\underline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\underline{x})$$

$$\leq f(\underline{x}). \qquad \blacksquare$$

Thursday: review session (i.e. office hours)

Exam: next Wed. (Dec 15), 8am. here (WEB L114)

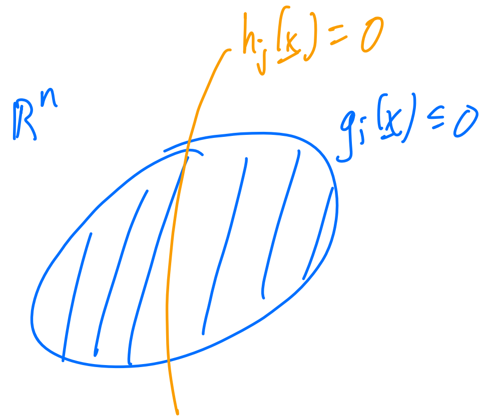
maybe?

Today: HW #6 due.

Review KKT conditions

$$\begin{aligned} \min \quad & f(\underline{x}) \\ \text{s.t.} \quad & g_i(\underline{x}) \le 0 \quad i=1,\dots m \\ & h_j(\underline{x}) = 0, \quad j=1,\dots p \end{aligned}$$

Constrained opt. problem



Define: $L(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(\underline{x}) + \sum_{i=1}^{m} \lambda_i g_i(\underline{x}) + \sum_{j=1}^{p} \mu_j h_j(\underline{x})$

$\underline{x}$ is a KKT point if:

$\nabla_{\underline{x}} L = \underline{0}$ (no feasible descent directions)

$\lambda_i g_i(\underline{x}) = 0$, $i=1,\dots m$ (complementary slackness)

$\underline{\lambda} \in \mathbb{R}^m_+$ $(\lambda_i \ge 0)$ (dual feasibility)

$$\left.\begin{aligned} \nabla_{\underline{\lambda}} L &\le \underline{0} \\ \nabla_{\underline{\mu}} L &= \underline{0} \end{aligned}\right\} \text{componentwise} \quad \text{(primal feasibility)}$$

KKT points aren't necessarily useful without "qualifications" on constraints.

If $\{\nabla g_i(\underline{x})\}_{i=1}^{m} \cup \{\nabla h_j(\underline{x})\}_{j=1}^{p}$ is a linearly independent set of vectors, then $\underline{x}$ is said to satisfy the linear independence constraint qualification (LICQ) condition.

Theorem (KKT necessity). Assume $\underline{x}^*$ satisfies the LICQ condition. $\underline{x}^*$ is a local minimum of

$$\begin{aligned} &\min f(\underline{x}) \\ &\text{s.t. } g_i(\underline{x}) \le 0, \quad i=1..m \\ &\quad\;\; h_j(\underline{x}) = 0, \quad j=1..p \end{aligned}$$

only if $\underline{x}^*$ ~~satisfies the LICQ condition and~~ is a KKT point

Ex (Constraint qualification is not optional)

$$\begin{aligned} &\min x \\ &\text{s.t. } -x^3 \le 0 \end{aligned} \qquad (x \in \mathbb{R}) \qquad \begin{aligned} &f(x) = x \\ &g(x) = -x^3 \end{aligned}$$

KKT conditions: $1 + \lambda(-3x^2) = 0 \quad (\nabla f + \lambda \nabla g = 0)$

$\lambda(-x^3) = 0 \quad (\lambda g(x) = 0)$

$\lambda \ge 0$

$x \ge 0$

There are no solutions : no KKT points
(but $x=0$ is a local min!)

Failure: $\nabla g(x=0) = (-3x^2)\big|_{x=0} = 0$ (not linearly independent)

(x is feasible + satisfies a constraint qualification
$\Longrightarrow$ x is "regular".)

Ex. (Reformulation can help)

$$\min \ x$$
$$\text{s.t.} \ -x \le 0 \qquad (x \in \mathbb{R})$$

$x=0$ is a KKT point.

(Here, KKT conditions are necessary.)

There are other kinds of constraint qualification.
E.g. linear constraint qualification (LCQ): g_i, h_j are all affine functions.

Thm (sufficiency of KKT)
Assume f is convex and g_j are all convex, and h_i are all affine: then if x^* is a KKT point, then it's a local minimum.

Ex: $\min x$
s.t. $-x^3 \leq 0$

$\Big\}$ f is convex
$g(x) = -x^3$ is concave (ok for sufficiency)
for $x \geq 0$.

There are no KKT points (but a local min exists!)