Recall the setting:

Let $f, g_1, \cdots, g_{m_1}, h_1, h_2, \cdots, h_{m_2} : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions. Consider

$$(Q) \qquad \begin{array}{ll} \inf & f(x) \\ \text{s.t.} & g_i(x) \le 0, \quad i = 1, \cdots, m_1, \quad (v_i) \\ & h_j(x) = 0, \quad j = 1, \cdots, m_2, \quad (w_j) \\ & x \in \mathbb{R}^n. \end{array} \quad \Big\} \; S: \text{feasible region}$$

Theorem (Karush-Kuhn-Tucker (KKT) necessary optimality condition)

Let $\bar{x} \in \mathbb{R}^n$ be a local minimum of (Q). Let

$$I = \{ i : g_i(\bar{x}) = 0 \}$$

be the active index set. Suppose that the Linear Independence Constraint Qualification (LICQ) holds:

$$\{\nabla g_i(\bar{x})\}_{i \in I} \cup \{\nabla h_j(\bar{x})\}_{j=1}^{m_2} \text{ are linearly independent.}$$

Then, $\exists$ multipliers $v_1, \cdots, v_{m_1}; w_1, \cdots, w_{m_2} \in \mathbb{R}$ s.t.

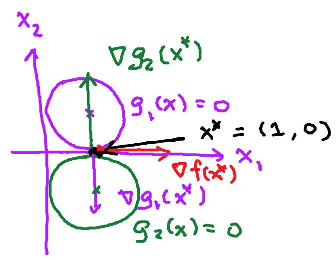
$$(*) \quad \begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m_1} v_i \nabla g_i(\bar{x}) + \sum_{j=1}^{m_2} w_j \nabla h_j(\bar{x}) = 0, \\ v_i \ge 0 \quad \forall i. \end{cases} \qquad \text{(dual feasibility)}$$

$$(**) \quad v_i \, g_i(\bar{x}) = 0 \quad \forall i. \qquad \text{(complementarity)}$$

Example (Importance of CQ)

Consider

$$\begin{array}{ll} \inf & f(x_1, x_2) = x_1 = e_1^T x \\ \text{s.t.} & g_1(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \le 0, \\ & g_2(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 1)^2 - 1 \le 0. \end{array}$$



Feasible set $= \{(1, 0)\} \Rightarrow$ optimal solution is $x^* = (1, 0)$.

The KKT conditions are given by

$$\underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\nabla f(x_1,x_2)} + \underbrace{v_1 \begin{bmatrix} 2(x_1-1) \\ 2(x_2-1) \end{bmatrix}}_{v_1 \nabla g_1(x_1,x_2)} + \underbrace{v_2 \begin{bmatrix} 2(x_1-1) \\ 2(x_2+1) \end{bmatrix}}_{v_2 \nabla g_2(x_1,x_2)} = 0,$$

$$v_1, v_2 \geq 0 \quad ; \quad v_i\, g_i(x_1,x_2) = 0 \quad \forall i.$$

Q: When $x = (1,0)$, can we find v_1, v_2 that solve the above?

No, since otherwise

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_1 \begin{bmatrix} 0 \\ -2 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 0,$$

which is impossible.

Other CQs?

Theorem: Suppose that $g_1, \dots, g_{m_1}$ are convex and $h_1, \dots, h_{m_2}$ are affine. (In particular, the feasible region is convex.) Suppose that the Slater CQ holds:

$$\exists\, x' \in S \text{ s.t. } g_i(x') < 0 \quad \forall i$$

Then, the KKT conditions $(*)$ and $(**)$ are necessary for optimality.

Remark: This applies to convex-constrained optimization problems.

Theorem: Suppose that $g_1, \dots, g_{m_1}$ are concave and $h_1, \dots, h_{m_2}$ are affine. Then, KKT conditions $(*)$ and $(**)$ are necessary for optimality.

Remark: This applies to linearly constrained optimization problems.

Example: Let $A \in S^n$ be given. Consider

$$(E) \qquad \begin{aligned} \min \quad & f(x) = x^T A x \\ \text{s.t.} \quad & h(x) = 1 - \|x\|_2^2 = 0 \end{aligned}$$

Observe that $\nabla h(x) = -2x \neq 0$ for any feasible x. Thus, if $\bar{x}$ is a local minimum of (E), then $\nabla h(\bar{x}) \neq 0 \Rightarrow$ LICQ holds.

The KKT optimality condition is

$$\underbrace{2Ax}_{\nabla f(x)} + \underbrace{w \cdot (-2x)}_{w \cdot \nabla h(x)} = 0 \quad \Rightarrow \quad Ax = wx.$$

$\Rightarrow$ (w, x) is an eigenpair of A.

eigenvalue   eigenvector

$\Rightarrow$ any optimal solution x^* to (E) is an eigenvector of A

Now, observe that

$$(x^*)^T A (x^*) \underset{Ax^* = wx^*}{=} (x^*)^T (wx^*) = w \, \|x^*\|_2^2 \underset{\|x^*\|_2^2 = 1}{=} w.$$

$\Rightarrow$ optimal value = smallest eigenvalue of A

optimal solution = corresponding eigenvector.

<u>Example</u>: Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ be given, Consider

$$
\text{(LP)} \qquad
\begin{array}{lll}
\min & c^T x & \longleftrightarrow \; f(x) = c^T x \\
\text{s.t.} & Ax = b & \; g_i(x) = \overset{-e_i^T x}{\overbrace{-x_i}} \le 0 \\
& x \geq 0 & \; h_j(x) = b_j - a_j^T x = 0
\end{array}
$$

Since this is a linearly constrained problem, the KKT conditions are necessary.

j^{th} col of A^T

$$\underbrace{c}_{\nabla f(x)} + \underbrace{\sum_{i=1}^{n} v_i (-e_i)}_{\sum_{i=1}^{n} v_i \nabla g_i(x)} + \underbrace{\sum_{j=1}^{m} w_j (-a_j)}_{\sum_{j=1}^{m} w_j \nabla h_j(x)} = 0, \qquad — (\Delta)$$

$$v_i \geq 0, \quad v_i x_i = 0 \quad \forall i$$

From (Δ), we have

$$\left.\begin{array}{l} c - v - A^T w = 0. \;\Rightarrow\; v = c - A^T w \\ \qquad\qquad\qquad\qquad\quad\; v \geq 0 \end{array}\right\} \Rightarrow \; c - A^T w \geq 0$$

$\Rightarrow$ KKT conditions become

$$A^T w \le c,$$

$$x_i (c - A^T w)_i = 0 \quad \forall i$$