## FJ conditions and Karush Kuhn Tucker (KKT ) conditions:

- Consider the following nonlinear programming problem (P) of the form,
- Minimize $f(\mathbf{x})$
- subject to $g_i(\mathbf{x}) \leq \mathbf{0}$, for $i = 1, \ldots, m$, $\quad \mathbf{x} \in \mathbb{R}^n$.
- All $f$,$g_i : \mathbb{R}^n \to \mathbb{R}$.
- $f$ and the constraint functions $g_i$ may not be linear functions.
- $S = Fea(P) = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, \text{ for } i = 1, ..., m\}$.

- The set **D** of **feasible directions** at $\mathbf{x}^* \in Fea(P)$ is given by,
  $\mathbf{D}_{\mathbf{x}^*} = \{\mathbf{d} \in \mathbb{R}^n : \exists \ \ c > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) \leq 0,$ for all $i = 1, ..., m,$ and for all $0 \leq t \leq c\}$.
- Let $I = \{i \in \{1, \ldots, m\} : g_i(\mathbf{x}^*) = 0\}$.
- Let $I^* = \{1, \ldots, m\} \setminus I = \{i \in \{1, \ldots, m\} : g_i(\mathbf{x}^*) < 0\}$.
- For all $i \in I^*$, we assume that $g_i$ is **continuous** at $\mathbf{x}^*$.
- For each $i \in I^*$ there exists $c_i > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for all $0 \leq t \leq c_i$,
- If $c = min_{i \in I^*}\{c_i\}$, then $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for all $0 \leq t \leq c$ and for all $i \in I^*$.

- For $i \in I$, $g_i$'s are assumed to be **continuously differentiable** at $\mathbf{x}^*$.
- If $\mathbf{d}$ satisfies the condition, $\nabla g_i(\mathbf{x}^*)\mathbf{d} < 0$ for all $i \in I$, then for each $i \in I$ there exists an $a_i > 0$ such that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for all $0 \leq t \leq a_i$.
- If $a = min_{i \in I}\{a_i\} > 0$, we get that $g_i(\mathbf{x}^* + t\mathbf{d}) < 0$ for all $0 \leq t \leq a$ and for all $i \in I$.
- If for all $i \in I^*$, $g_i$'s are assumed to be **continuous** at $\mathbf{x}^*$ and for all $i \in I$, $g_i$'s are assumed to be **continuously differentiable** at $\mathbf{x}^*$ then
  $\mathbf{G}_0 \subseteq \mathbf{D}_{\mathbf{x}^*}$,
  where $\mathbf{G}_0 = \{\mathbf{d} \in \mathbb{R}^n : \nabla g_i(\mathbf{x}^*)\mathbf{d} < 0 \text{ for all } i \in I\}$.

- If $\mathbf{x}^*$ is a local **minimum**, then $\mathbf{F}_0 \cap \mathbf{D}_{\mathbf{x}^*} = \phi$,where $\mathbf{D}_{\mathbf{x}^*}$ is the set of feasible directions at $\mathbf{x}^*$ and $\mathbf{F}_0 = \{\mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)\mathbf{d} < 0\}$.
- Since $\mathbf{G}_0 \subseteq \mathbf{D}_{\mathbf{x}^*}$ if $\mathbf{x}^*$ is a **local minimum** implies $\mathbf{F}_0 \cap \ \mathbf{G}_0 = \phi$.
- **Theorem 7:** (**FJ necessary conditions**) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a **continuously differentiable** function.
  Consider the problem **P** of minimizing $f$ subject to the conditions $g_i(\mathbf{x}) \leq 0, i = 1, ..., m$, where $g_i : \mathbb{R}^n \to \mathbb{R}$ for all $i$.
  Let $S = Fea(P) = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, i = 1, ..., m\}$ and $\mathbf{x}^* \in S$.
  For all $i \in I^*$, $g_i$'s are **continuous** at $\mathbf{x}^*$
  and for all $i \in I$, $g_i$'s are **continuously differentiable** at $\mathbf{x}^*$.
  If $\mathbf{x}^*$ is a **local minimum** of $f$ over $S$
  there exists **non negative** constants, $u_0, u_i, i \in I$, **not** all zeros such that
  $u_0 \nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$. **(*)**

- If all the $g_i$'s for $i = 1, \ldots, m$, are **continuously differentiable** at $\mathbf{x}^*$ then the above condition reduces to
- $u_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$
and $u_i g_i(\mathbf{x}^*) = 0$, for all $i = 1, ..., m$. **(**)**
- $\mathbf{x}^* \in S$ is called the **primal feasibility condition**,
- **(*)** together with nonnegativity of the $u_i$'s, is called the **dual feasibility condition.**
- $u_i g_i(\mathbf{x}^*) = 0$, for all $i = 1, ..., m$, is called the **complementary slackness condition.**
- The above conditions are called the **FJ (Fritz John)** conditions and the point $(\mathbf{x}^*, \mathbf{u})$ ( or $\mathbf{x}^*$) is called a **Fritz John**, or an **FJ**, point.
- $\mathbf{x}^*$ is said to satisfy **KKT** condition if there exists non negative constants $u_i$, $i \in I$, such that
$\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$. **(***)**

- The above conditions given by (***) are called **KKT (Karush, Kuhn, Tucker ) conditions.**
  Any $(\mathbf{x}^*, \mathbf{u})$ (or $\mathbf{x}^*$) which satisfies the **Karush Kuhn Tucker (KKT)** conditions is called a **KKT** point.
- **Theorem 8:** (**KKT necessary conditions**) If in addition to the conditions assumed for $f$ and $g_i$'s, as in the previous theorem, if $\nabla g_i(\mathbf{x}^*)$'s for $i \in I$ are **LI**
  and $\mathbf{x}^*$ is a local minimum of $f$ then there exists nonnegative constants $u_i$, $i \in I$, such that
  $\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$. **(****)**
- **Remark:** Hence from **Theorem 8** it is clear that if $(\mathbf{x}^*, \mathbf{u})$ ( or $\mathbf{x}^*$) is an **FJ** point and if $\nabla g_i(\mathbf{x}^*)$'s for $i \in I$ are **LI** then $(\mathbf{x}^*, \mathbf{u})$ ( or $\mathbf{x}^*$) is a **KKT** point.

- **Remark:** If $\mathbf{x}^*$ is a **Karush Kuhn Tucker (KKT)** point then it is also an **FJ** point.
- **Remark: KKT** conditions says that under the assumptions of the above theorem, if $\mathbf{x}^*$ is a **local minimum**, then $-\nabla f(\mathbf{x}^*)$ lies in the **cone generated** by the $\nabla g_i(\mathbf{x}^*)$'s, $i \in I$.
- **Gordon's Theorem:** Exactly one of the following two systems has a solution:
  $\mathbf{u} \neq \mathbf{0}$, $\mathbf{u} \geq \mathbf{0}$, $\mathbf{u}^T A = \mathbf{0}$ **(1)**
  $A\mathbf{y} > \mathbf{0}$ **(2)**

- **Example 1:** $f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2$
  subject to
  $x_1^2 + x_2^2 \leq 5.$
  $x_1 + 2x_2 \leq 4.$
  $-x_1 \leq 0$
  $-x_2 \leq 0.$
- $f$ takes its minimum value at $[2, 1]^T$.
- $g_1(\mathbf{x}) = x_1^2 + x_2^2 - 5$,
  $g_2(\mathbf{x}) = x_1 + 2x_2 - 4$, $g_3(\mathbf{x}) = -x_1$ and $g_4(\mathbf{x}) = -x_2$.
- At $\mathbf{x}^* = [2, 1]^T$ the binding constraints are $g_1$ and $g_2$.
- $\nabla g_1(\mathbf{x}^*) = [4, 2]$, $\nabla g_2(\mathbf{x}^*) = [1, 2]$, $\nabla f(\mathbf{x}^*) = [-2, -2]$.
  $\nabla g_1(\mathbf{x}^*)$ and $\nabla g_2(\mathbf{x}^*)$ are LI.
- Take $u_1 = \frac{2}{6}$, $u_2 = \frac{2}{3}$, then
- $\nabla f(\mathbf{x}^*) + \sum_{i=1}^{2} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.
  Hence $[2, 1]^T$ satisfies the KKT condition.

- **Example 2:** $f(x_1, x_2) = -x_1$.
  subject to
  $x_2 - (1 - x_1)^3 \leq 0$
  $-x_2 \leq 0$
- $\mathbf{x}^* = [1, 0]^T$ is a **local minimum.**
- At $[1, 0]^T$ both the constraints are binding.
- $\nabla f(\mathbf{x}^*) = [-1, 0]$, $\nabla g_1(\mathbf{x}^*) = [0, 1]$ and $\nabla g_2(\mathbf{x}^*) = [0, -1]$.
- Take $u_0 = 0, u_1 = 1, u_2 = 1$, then $u_0 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{2} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.
- Since $-\nabla f(\mathbf{x}^*)$ does not lie in the cone generated by the $\nabla g_i(\mathbf{x}^*)$'s, $i = 1, 2$,
  $[1, 0]^T$ is an **FJ point** but **not** a **KKT point.**
- So **KKT** condition is **not** a necessary condition for a **local minimum**, although FJ conditions are **necessary conditions** for a **local minimum.**

- **Theorem 8:** (**KKT necessary conditions**) If in addition to the conditions assumed for $f$ and $g_i$'s, as in the previous theorem, if $G_0 \neq \phi$
  and $\mathbf{x}^*$ is a local minimum of $f$ then there exists nonnegative constants $u_i$, $i \in I$, such that
  $\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$. **(**)**
- **Example 2:** Minimize $-x_1$.
  subject to
  $-x_2 - x_1 \leq 0$
  $-x_2 \leq 0$
  $-x_1 \leq 0$
- $\nabla g_i(\mathbf{x}^*)$'s, $i \in I$ are **LD** at $\mathbf{x}^* = [0, 0]^T$ but $G_0 \neq \phi$.
- **Remark:** If $\nabla g_i(\mathbf{x}^*)$'s, $i \in I$ are **LI** then $G_0 \neq \phi$ but the converse is **not** true.

- **Theorem 9:(KKT sufficient conditions)** Let $f : \mathbb{R}^n \to \mathbb{R}$ be **convex** and **continuously differentiable.**
  Consider the problem of **minimizing** $f$ subject to the conditions $g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m$,
  where $g_i : \mathbb{R}^n \to \mathbb{R}$ for all $i$.
  Let $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\}$ and $\mathbf{x}^* \in S$. For all $i \in I^*$, assume that $g_i$ is **continuous** at $\mathbf{x}^*$
  and for all $i \in I$, $g_i$'s are assumed to be **continuously differentiable** at $\mathbf{x}^*$.
  Let all the $g_i$'s be **convex functions**, so that
  $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, i = 1, 2, \ldots, m\}$ is **convex.** Then
  $\mathbf{x}^*$ is a **global minimum** of $f$ over $S$ if there exists non negative constants, $u_i, i \in I$ such that
  $\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$. **(*)**

- **Example 1 revisited:** $f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2$
  subject to
  $x_1^2 + x_2^2 \leq 5.$
  $x_1 + 2x_2 \leq 4.$
  $-x_1 \leq 0$
  $-x_2 \leq 0.$
- Check that $f, g_i$ for all $i$ are convex functions.
- $\mathbf{x}^* = [2, 1]^T$ is a KKT point.
- $\mathbf{x}^* = [2, 1]^T$ is the global minimum of $f$.

- **Example 3:** $f(x) = -x^2$ for $x \leq 0$
  $= x^2$ for $x \geq 0$.
- For $x^* = 0$, $\nabla f(x^*) = 0$, hence $x^* = 0$ satisfies the **KKT** conditions but $x^*$ not be a local minimum point.
- $f$ is clearly **not** convex.
- Any $x^*$ for which $\nabla f(x^*) = 0$ is a KKT point but $x^*$ may not be a point of local minimum.
- $f(x_1, x_2) = x_1$
  subject to
  $(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1$
  $(x_1 - 1)^2 + (x_2 + 1)^2 \leq 1$
  $x^* = [1, 0]^T$ is the **global minimum** of f but not a KKT point.

- **Exercise:** The KKT conditions for the linear programming problem (LPP):
  Min $\mathbf{c}^T\mathbf{x}$
  subject to, $A_{m\times n}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$ are given by:
- $\mathbf{c}^T - \mathbf{y}^T A - \mathbf{v}^T = \mathbf{0}$, or $\mathbf{c}^T - \mathbf{y}^T A = \mathbf{v}^T$
  and $v_i x_i = 0$ for all $i = 1, ...., n$.
  where $\mathbf{v} = [v_1, \ldots, v_n]^T$ is a non negative vector,
  and $\mathbf{y} = [y_1, \ldots, y_m]^T$ is unrestricted in sign.
- The above conditions are equivalent to the **dual feasibility** and the **complementary slackness** conditions for the above LPP.
- **Conclusion:** $\mathbf{x}^* \in Fea(P)$ is **optimal** for (P) if and only if $\mathbf{x}^*$ is a **KKT point** of (P).
- **Exercise:** What are the FJ points of the above problem?
- **Exercise:** What are the KKT conditions for the following linear programming problem
  Min $\mathbf{c}^T\mathbf{x}$
  subject to, $A\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}$.

- **Theorem 10:** (**FJ necessary conditions**) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a **continuously differentiable** function.
Consider the problem **P** of minimizing $f$ subject to the conditions
$g_i(\mathbf{x}) \leq 0, i = 1, ..., m$, where $g_i : \mathbb{R}^n \to \mathbb{R}$ for all $i$,
$h_j(\mathbf{x}) = 0, j = 1, ..., l$, where $h_j : \mathbb{R}^n \to \mathbb{R}$ for all $j$.
Let $S = Fea(P)$ and $\mathbf{x}^* \in S$.
For all $i \in I^*$, $g_i$'s are **continuous** at $\mathbf{x}^*$,
for all $i \in I$, $g_i$'s are **continuously differentiable** at $\mathbf{x}^*$,
and for all $j$, $h_j$'s are **continuously differentiable** at $\mathbf{x}^*$.
Then if $\mathbf{x}^*$ is a **local minimum** of $f$ over $S$
there exists $u_0, u_i, i \in I$, **non negative** constants, and $v_j$ constants (**unrestricted in sign**), not all zeros, such that
$u_0 \nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{l} v_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$. **(*)**

- **Example 5:** Minimize $x_1 + x_2$
  subject to
  $x_2 - x_1 = 0$.
- Check that no point of the above problem satisfies the **FJ** conditions (as in Theorem 10).
- **Example 6:** Minimize $(x_1 - 1)^2 + x_2$
  subject to
  $x_2 - x_1 = 1$
  $x_1 + x_2 \leq 2$.
- Check that $\mathbf{x}^* = [\frac{1}{2}, \frac{3}{2}]^T$ is a **KKT** point with $u = 0, v = -1$ (as in Theorem 10).

- **Theorem 11:** (**KKT necessary conditions**) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a **continuously differentiable** function. Consider the problem **P** of minimizing $f$ subject to the conditions
  $g_i(\mathbf{x}) \leq 0, i = 1, ..., m$, where $g_i : \mathbb{R}^n \to \mathbb{R}$ for all $i$,
  $h_j(\mathbf{x}) = 0, j = 1, ..., l$, where $h_j : \mathbb{R}^n \to \mathbb{R}$ for all $j$.
  Let $S = Fea(P)$ and $\mathbf{x}^* \in S$.
  For all $i \in I^*$, $g_i$'s are **continuous** at $\mathbf{x}^*$,
  for all $i \in I$, $g_i$'s are **continuously differentiable** at $\mathbf{x}^*$,
  and for all $j$, $h_j$'s are **continuously differentiable** at $\mathbf{x}^*$.
  Let $\nabla g_i(\mathbf{x}^*)$'s, $\nabla h_j(\mathbf{x}^*)$ for $i \in I, j \in \{1, \ldots, l\}$ be **LI**. Then if $\mathbf{x}^*$ is a **local minimum** of $f$ over $S$
  there exists $u_i, i \in I$, **non negative** constants, and $v_j$ constants (**unrestricted in sign**) such that
  $\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{l} v_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$. **(*)**

- **Example 7:** Minimize $f(x_1, x_2) = x_1^3 - x_1 + x_2$,
  subject to
  $(x_1 - 1)^2 + 2x_2 \leq 1$,
  $x_1 - x_2 = 0$.
- Check that at $\mathbf{x}^* = [0, 0]^T$, the first constraint is binding.
- But $\nabla g_1(\mathbf{x}^*)) = 2[-1, 1]^T$, and $\nabla h_1(\mathbf{x}^*) = [1, -1]^T$.
- Since $\{\nabla g_1(\mathbf{x}^*)), \nabla h_1(\mathbf{x}^*)\}$ is **LD**, Theorem 11 is not applicable.
- Check that $[0, 0]^T$ satisfies the KKT conditions as given in **(3)**, immediately after Theorem 7.
- Since $f$, $g_1, g_2, g_3$ is convex, where $g_2(x_1, x_2) = x_1 - x_2$ and $g_3(x_1, x_2) = -x_1 + x_2$,
- $[0, 0]^T$ is the global minimum of $f$ in the given feasible region.