Optimality conditions for constrained optimization problems : $\begin{cases} \min & f(x) \\ \text{s.t.} & g(x) \leq 0 \\ & h(x) = 0 \end{cases}$ (P)

Fritz John Necessary Conditions : Let $\bar{x}$ be a feasible point of (P).

If $\bar{x}$ is a local minimum of (P), then there exists $(u_0, u, v) \neq 0$ such that

$$u_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) + \sum_{i=1}^{\ell} v_i \nabla h_i(\bar{x}) = 0$$

$$u_0, u \geq 0$$

$$u_i g_i(\bar{x}) = 0 \quad , i = 1, \dots, m.$$

Definitions: Let $\bar{x}$ be a feasible point of (P).

$F_0 := \{ d \in \mathbb{R}^n : (\nabla f)(\bar{x})^T d < 0 \}$ is the set of descent directions.

$I := \{ i : g_i(\bar{x}) = 0 \}$ is the set of indices of active constraints (binding constraints)

$G_0 := \{ d \in \mathbb{R}^n : \nabla g_i(\bar{x})^T d < 0 \text{ for all } i \in I \}$ is the set of "inward" pointing directions for the active constraints at $\bar{x}$.

$H_0 := \{ d \in \mathbb{R}^n : \nabla h_i(\bar{x})^T d = 0 \text{ for all } i = 1, \dots, \ell \}$ is the set of tangent directions for the equality constraints at $\bar{x}$.

Theorem 1 (Geometric First-Order Necessary Condition):

If $\bar{x}$ is a local minimum of (P) : $\begin{cases} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \ i = 1, \dots, m \\ & h_i(x) = 0, \ i = 1, \dots, \ell \end{cases}$ and either

(i) $h(x) = Ax - b$, or

(ii) $\nabla h_i(\bar{x})$, $i = 1, \dots, \ell$ are linearly independent,

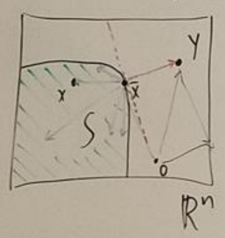
then $F_0 \cap G_0 \cap H_0 = \emptyset$.

"if a point $\bar{x}$ is locally optimal, there is no direction d which is both a feasible direction ($g(\bar{x} + \lambda d) \leq 0$, $h(\bar{x} + \lambda d) \approx 0$ for small λ) and an improving direction w.r.t. the cost ($f(\bar{x} + \lambda d) < f(\bar{x})$ for small λ)."

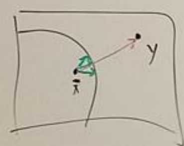
Goal next: Rephrase $F_0 \cap G_0 \cap H_0 = \emptyset$ into "Computable" algebraic conditions about $\nabla f, \nabla g, \nabla h$
$\rightsquigarrow$ Fritz John necessary conditions.

To do so, we review some fundamental results from Mathematical Optimization Theory:

Proposition 1: Let S be a nonempty closed convex set in $\mathbb{R}^n$, and $y \notin S$ ($y \in \mathbb{R}^n \setminus S$) ($y \in S^c$)
Then there exists a unique point $\bar{x} \in S$ of minimum distance to y.
Furthermore, $\bar{x}$ is the minimizing point if and only if



$$(y-\bar{x})^T(x-\bar{x}) \le 0 \quad \forall x \in S. \qquad (1)$$



Proof: Let $f(x) = \|x-y\|^2$. Let $\hat{x}$ be an arbitrary point in S, and define $R := \|\hat{x}-y\|$.
Then consider $\hat{S} = S \cap \underbrace{\overline{B_R(y)}}_{=\{x:\ \|x-y\| \le R\}}$ to obtain a compact search region (closed & bounded)

(Weierstrass extreme value theorem: Continuous functions attain their minimum over compact sets)

$\Rightarrow \exists\, \bar{x} \in \hat{S} \subset S$ that minimizes f over S (existence ✓)

uniqueness: $f(x) = \|x-y\|^2$ is strictly convex ($\hat{S}$ is convex) $\Rightarrow \bar{x}$ is unique.

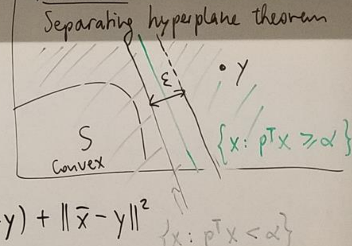
Next we show that $\bar{x} \in S$ closest to y $\overset{\Rightarrow}{\underset{\Leftarrow}{\text{if and only if}}}$ (1) holds.

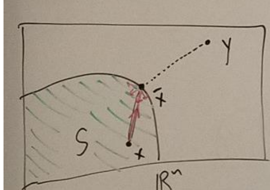
$x^T y = y^T x$

"$\Leftarrow$": Note that $\forall x \in S$: $\|x-y\|^2 = \|(x-\bar{x})-(y-\bar{x})\|^2 = \underbrace{\|x-\bar{x}\|^2}_{\ge 0} + \|y-\bar{x}\|^2 - 2\underbrace{(x-\bar{x})^T(y-\bar{x})}_{\le 0 \text{ from (1)!}} \ge \|\bar{x}-y\|^2$

"$\Rightarrow$": Assume $\bar{x} \in S$ that minimizes the distance to y.
For any $x \in S$, $\lambda x + (1-\lambda)\bar{x} \in S \quad \forall \lambda \in [0,1]$. Therefore
$\|\lambda x + (1-\lambda)\bar{x} - y\| \ge \|\bar{x}-y\|$. Thus:

PREVIEW:
Separating hyperplane theorem





$$\|\bar{x}-y\|^2 \le \|\lambda x + (1-\lambda)\bar{x} - y\|^2$$
$$= \|\lambda(x-\bar{x}) + (\bar{x}-y)\|^2 = \lambda^2\|x-\bar{x}\|^2 + 2\lambda(x-\bar{x})^T(\bar{x}-y) + \|\bar{x}-y\|^2$$

which when rearranged yields: $\lambda^2 \|x-\bar{x}\|^2 \ge 2\lambda (y-\bar{x})^T(x-\bar{x})$.

Focusing on $\lambda > 0$ and dividing by λ: $\lambda\|x-\bar{x}\|^2 \ge 2(y-\bar{x})^T(x-\bar{x}) \quad \Big|\ \forall \lambda > 0$

Letting $\lambda \to 0$: $\quad 0 \ge (y-\bar{x})^T(x-\bar{x})$. $\blacksquare$