Lecture 2: Optimality conditions for constrained optimization – KKT conditions

- Brief recap; optimization problems and convexity
- Motivating examples for KKT conditions
- KKT conditions

Reference: Chapter 12.1, 12.2 in N&W

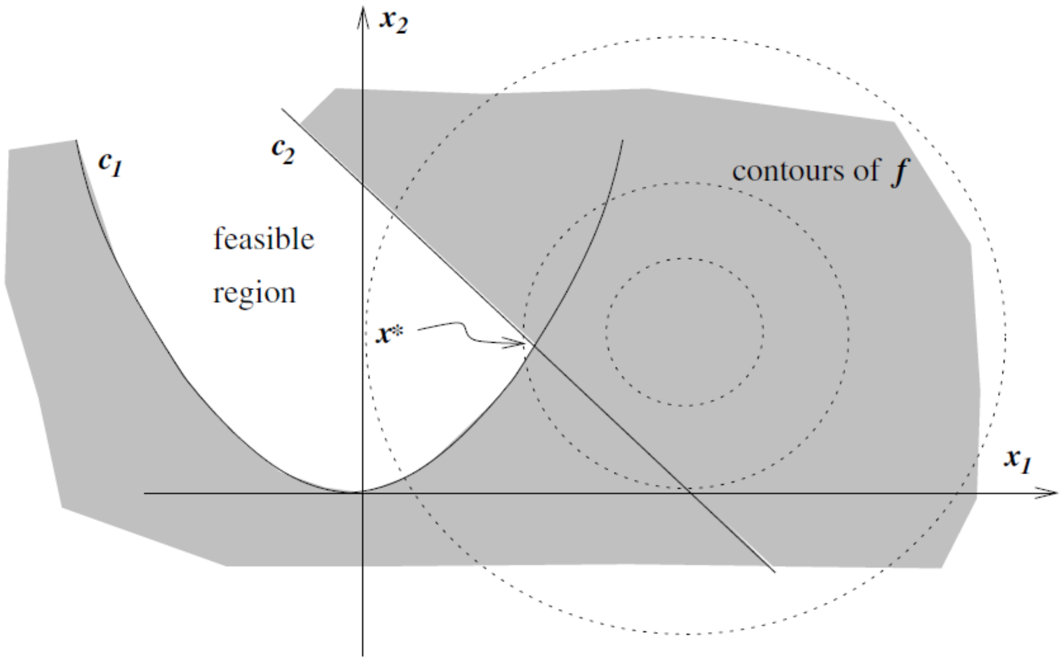
Administrative information

- Check announcements on Blackboard regularly.
- Assignments posted every Monday, due two weeks later (Wednesdays). Deliver on Blackboard or to room D238.
- The lab will start in week 5 and end in week 11. Make groups of 2 or 3 people. More information on Blackboard.
- Must complete 7/10 assignments and the lab to attend the exam.
  - Exercise 0 does not count

General optimization problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{array}{ll} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}. \end{array}$$

- Example: $\min\ (x_1 - 2)^2 + (x_2 - 1)^2 \quad \text{subject to} \quad \begin{array}{l} x_1^2 - x_2 \leq 0, \\ x_1 + x_2 \leq 2. \end{array}$



- What if we add equality-constraint $x_1 = 0$?

Definitions

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases} \qquad \text{(P)}$$

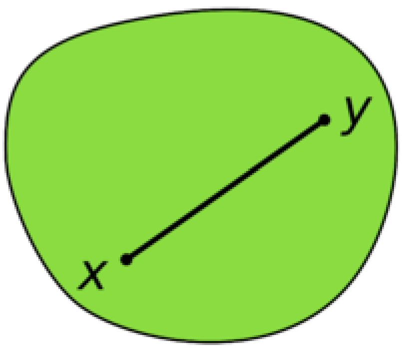
- Feasible set: $\Omega = \{x \in \mathbb{R}^n \mid c_i(x) = 0,\ i \in \mathcal{E};\ c_i(x) \geq 0,\ i \in \mathcal{I}\}$

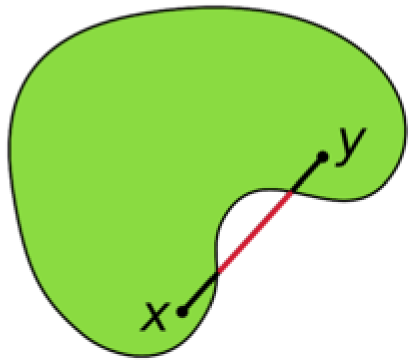
- A vector x^* is a *global solution* to (P) if $x^* \in \Omega$ and $f(x) \geq f(x^*)$ for $x \in \Omega$.

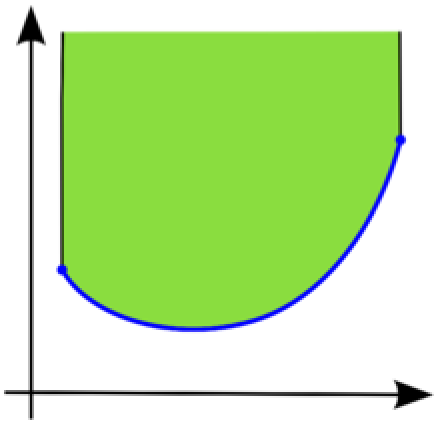
- A vector x^* is a *local solution* to (P) if $x^* \in \Omega$ and there is a neighborhood $\mathcal{N}$ of x^* such that $f(x) \geq f(x^*)$ for $x \in \mathcal{N} \cap \Omega$.

- A vector x^* is a *strict local solution* to (P) if $x^* \in \Omega$ and there is a neighborhood $\mathcal{N}$ of x^* such that $f(x) > f(x^*)$ for $x \in \mathcal{N} \cap \Omega$ with $x \neq x^*$.

Convexity







wikipedia.org

If the line segment between any two points within a set is inside the set, the set is **convex**.

A function is **convex** if the epigraph is a convex set.

- A convex optimization problem: Both *f(x)* and the feasible set convex
- Convex optimization problems are preferable!
  - For convex optimization problems, every local minimum is also a global minimum. **Sufficient to search for a local minimum!** Which is much easier than searching for global minimum.
  - For many convex optimization problems, it is easy to find derivatives, exploit structure, etc. making them easier to solve.
  - They typically have “guaranteed complexity”.

Convexity, cont'd

- When is

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases}$$

  convex?

- Condition:
  - $f(x)$ is a convex function:

$$\forall x, y \in \Omega,\ \forall \alpha \in [0,1]: \quad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

  - The feasible set

$$\Omega = \{x \in \mathbb{R}^n | c_i(x) = 0,\ i \in \mathcal{E},\ c_i(x) \geq 0,\ i \in \mathcal{I}\}$$

    is convex:

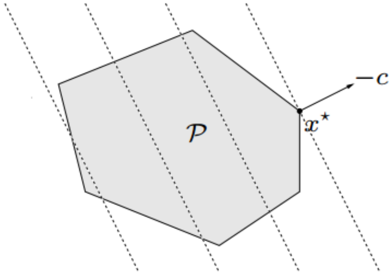
$$\forall x, y \in \Omega,\ \forall \alpha \in [0,1]: \quad \alpha x + (1-\alpha)y \in \Omega$$

- When is the feasible set convex?
  - $c_i(x),\ i \in \mathcal{E}$ are linear
  - $c_i(x),\ i \in \mathcal{I}$ are concave

Types of constrained optimization problems

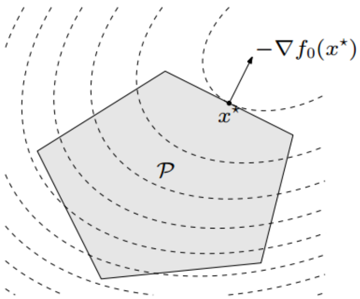
- Linear programming
  - Convex problem
  - Feasible set polyhedron

$$\begin{array}{ll} \text{minimize} & c^{\mathsf{T}}x \\ \text{subject to} & Ax \leq b \\ & Cx = d \end{array}$$



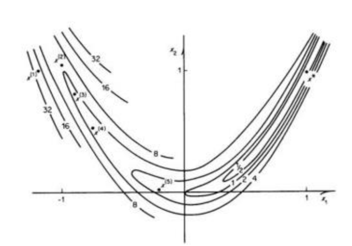
- Quadratic programming
  - Convex problem if $P \geq 0$
  - Feasible set polyhedron

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^{\mathsf{T}}Px + q^{\mathsf{T}}x \\ \text{subject to} & Ax \leq b \\ & Cx = d \end{array}$$



- Nonlinear programming
  - In general non-convex!

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) = 0 \\ & h(x) \geq 0 \end{array}$$



$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}. \end{cases}$$

Necessary conditions for unconstrained optimization problems

Theorem 2.2 (First-Order Necessary Conditions).

If x^* is a local minimizer and f is continuously differentiable in an open neighborhood of x^*, then $\nabla f(x^*) = 0$.

- What about constrained problems?

KKT conditions

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}. \end{cases}$$

- Lagrangian

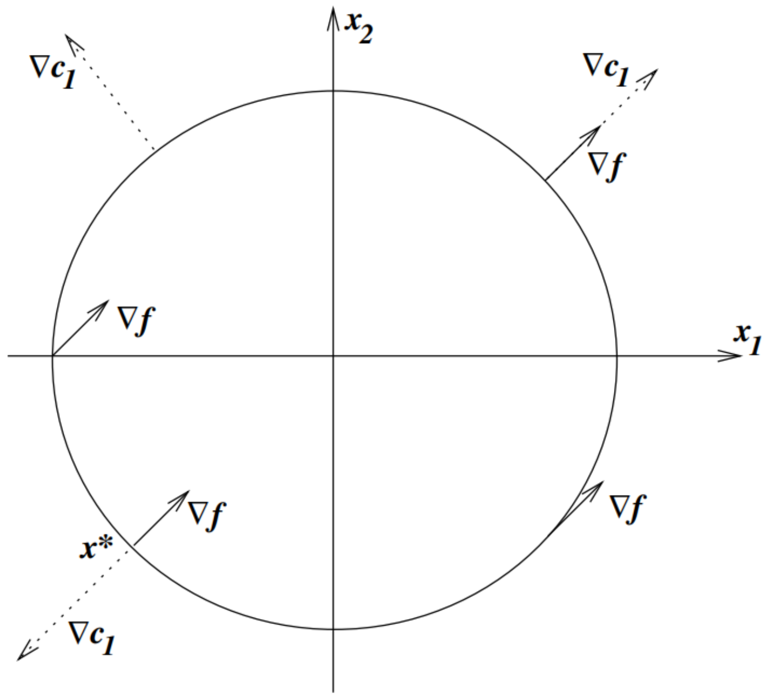
$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$$

- KKT conditions:

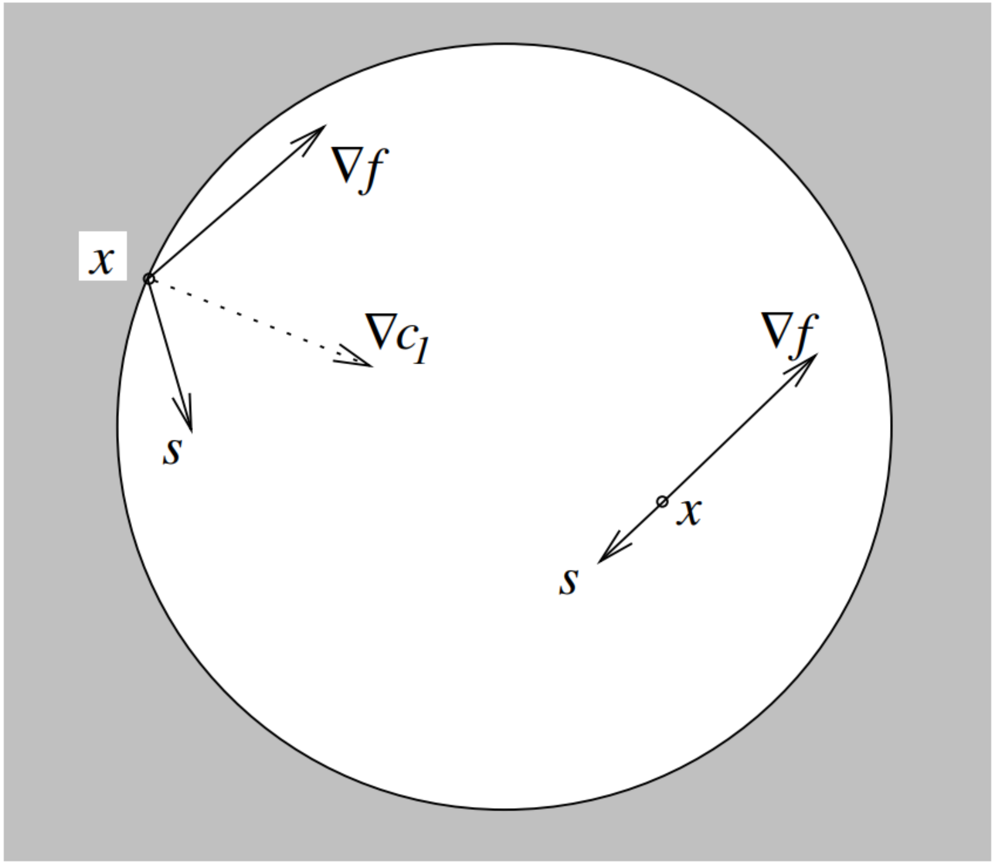
$$\begin{aligned}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\
c_i(x^*) &= 0, \quad \forall i \in \mathcal{E}, \\
c_i(x^*) &\geq 0, \quad \forall i \in \mathcal{I}, \\
\lambda_i^* &\geq 0, \quad \forall i \in \mathcal{I}, \\
\lambda_i^* c_i(x^*) &= 0, \quad \forall i \in \mathcal{E} \cup \mathcal{I}.
\end{aligned}$$

Example 12.1

$$\min\ x_1 + x_2 \qquad \text{s.t.} \qquad x_1^2 + x_2^2 - 2 = 0$$



Example 12.2



Active set

The active set $\mathcal{A}(x)$ at any feasible point x consists of the equality constraint indices from $\mathcal{E}$ together with the indices of the inequality constraints i for which $c_i(x) = 0$. That is,

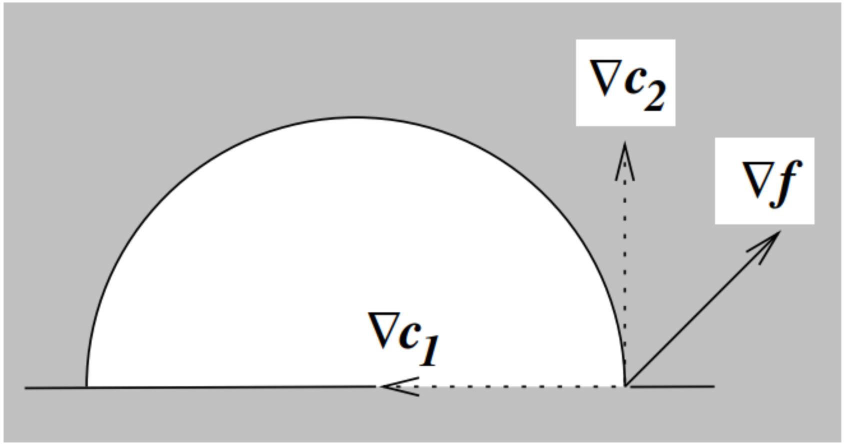
$$\mathcal{A}(x) = \mathcal{E} \cup \left\{ i \in \mathcal{I} \middle| c_i(x) = 0 \right\}$$

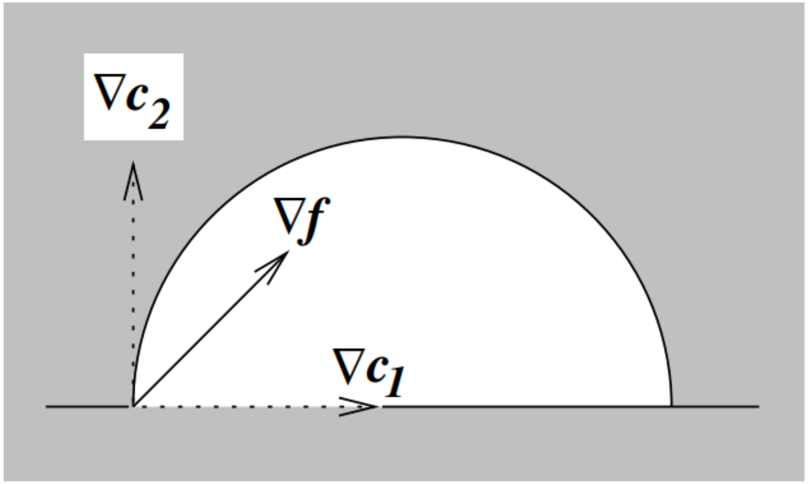
Set of feasible directions

Given a feasible point x and the active constraint set $\mathcal{A}(x)$, the set of linearized feasible directions $\mathcal{F}(x)$ is

$$\mathcal{F}(x) = \left\{ d \;\middle|\; \begin{array}{ll} d^\top \nabla c_i(x) = 0, & \text{for all } i \in \mathcal{E}, \\ d^\top \nabla c_i(x) \geq 0, & \text{for all } i \in \mathcal{A}(x) \cap \mathcal{I} \end{array} \right\}$$

Example 12.3





$$\min_{x \in \mathbb{R}^n} f(x) \qquad \text{subject to} \begin{cases} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}, \end{cases} \tag{12.1}$$

Theorem 12.1 (First-Order Necessary Conditions).

Suppose that x^* is a local solution of (12.1), that the functions f and c_i in (12.1) are continuously differentiable, and that the LICQ holds at x^*. Then there is a Lagrange multiplier vector λ^*, with components λ_i^*, $i \in \mathcal{E} \cup \mathcal{I}$, such that the following conditions are satisfied at (x^*, λ^*)

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \tag{12.34a}$$

$$c_i(x^*) = 0, \qquad \text{for all } i \in \mathcal{E}, \tag{12.34b}$$

$$c_i(x^*) \geq 0, \qquad \text{for all } i \in \mathcal{I}, \tag{12.34c}$$

$$\lambda_i^* \geq 0, \qquad \text{for all } i \in \mathcal{I}, \tag{12.34d}$$

$$\lambda_i^* c_i(x^*) = 0, \qquad \text{for all } i \in \mathcal{E} \cup \mathcal{I}. \tag{12.34e}$$

LICQ

Given the point x and the active set $\mathcal{A}(x)$, we say that the linear independence constraint qualification (LICQ) holds if the set of active constraint gradients $\{\nabla c_i(x),\ i \in \mathcal{A}(x)\}$ is linearly independent.

Why are KKT-conditions so important?

- Most algorithms for constrained optimization look for candidate solutions that fulfill KKT conditions
  - These are iterative algorithms that stop when KKT conditions fulfilled

- And:
  - When faced with a problem that you don't know how to handle, write down the optimality conditions
  - Often you can learn a lot about a problem, by examining the properties of its optimal solutions.