12 Theory of Constrained Optimization
General form

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

- E, I are index sets for equality and inequality constraints
- Feasible set \( \Omega = \{ x | c_i(x) = 0, i \in E; c_i(x) \geq 0, i \in I \} \)

Outline
- Equality and inequality constraints
- Lagrange multipliers
- Linear independent constraint qualification
- First/second order optimality conditions
- Duality
A single equality constraint

- An example \( \min_x f(x) = x_1 + x_2 \)
  subject to \( c_1 = x_1^2 + x_2^2 = 2 \)

\[
\nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \nabla c_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}
\]

- The optimal solution is at \( \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \)

- Optimal condition: \( \nabla f(x^*) = \lambda^* \nabla c_1(x^*) \)
  \( \lambda^* = -1/2 \) in the example
A single inequality constraint

- **Example:** \( \min_x f(x) = x_1 + x_2 \)
  subject to \( c_1 : 2 - x_1^2 - x_2^2 \geq 0 \)

- **Case 1:** the solution is inside \( c_1 \).
  - Unconstrained optimization \( \nabla f(x^*) = 0 \)

- **Case 2:** the solution is on the boundary of \( c_1 \).
  - Equality constraint
    \[
    \begin{align*}
    c_1(x^*) &= 0 \\
    \nabla f(x^*) &= \lambda^* \nabla c_1(x^*)
    \end{align*}
    \]
  - Complementarity condition: \( \lambda^* c_1(x^*) = 0 \)
    - Let \( \lambda^* = 0 \) in case 1.
Lagrangian function

- Define $L(x, \lambda) = f(x) - \lambda c_1(x)$
  - $\nabla_\lambda L(x, \lambda) = -c_1(x)$
  - $\nabla_x L(x, \lambda) = \nabla f(x) - \lambda \nabla c_1(x)$

- The optimality conditions of inequality constraint
  \[
  \begin{align*}
  c_1(x^*) &= 0 \\
  \nabla f(x^*) &= \lambda^* \nabla c_1(x^*)
  \end{align*}
  \]
  which equal to $\nabla_\lambda L(x, \lambda) = 0$ and $\nabla_x L(x, \lambda) = 0$.

- $L(x, \lambda)$ is called the Lagrangian function; $\lambda$ is called the Lagrangian multiplier.
Lagrangian multiplier

- At optimal solution $x^*$, $\nabla f(x^*) = \lambda_1 \nabla c_1(x^*)$
- If $c_1(.)$ changes $\delta$ unit, $f(.)$ changes $\approx \lambda_1 \delta$ unit.
  - $\lambda_1$ is the change of $f$ per unit change of $c_1$.
  - $\lambda_1$ means the sensitivity of $f$ to $c_i$.
  - $\lambda_1$ is called the *shadow price* or the *dual variable*. 
Constraint qualification

- **Active set** \( \mathbf{A}(x) = \mathbf{E} \cup \{ i \in \mathbf{I} \mid c_i(x) = 0 \} \)
- Linear independent constraint qualification (LICO)
  - The gradients of constraints in the active set \( \{ \nabla c_i(x) \mid i \in \mathbf{A}(x) \} \) are linearly independent
- A point is called regular if it satisfies LICO.
  - Other constraint qualifications may be used.

\[
\begin{align*}
  c_1(x) & = 1 - x_1^2 - (x_2 - 1)^2 \geq 0 \\
  c_2(x) & = -x_2 \geq 0 \\
  x & = (0, 0)
\end{align*}
\]
First order conditions

- A regular point that is a minimizer of the unconstrained problem must satisfy the KKT condition (Karush-Kuhn-Tucker)

\[ \nabla_x L(x^*, \lambda^*) = 0 \]

\[ c_i(x^*) = 0 \quad \text{for all } i \in \mathbf{E} \]

\[ c_i(x^*) \geq 0 \quad \text{for all } i \in \mathbf{I} \]

\[ \lambda^* \geq 0 \quad \text{for all } i \in \mathbf{I} \]

\[ \lambda^* c_i(x^*) = 0 \quad \text{for all } i \]

- The last one is called *complementarity* condition
Second order conditions

• The critical cone $C(x^*, \lambda^*)$ is a set of vectors that

\[ w \in C(x^*, \lambda^*) \iff \begin{cases} \nabla c_i(x^*)^T w = 0 & i \in \mathbf{E} \\ \nabla c_i(x^*)^T w = 0 & i \in \mathbf{I} \cap \mathbf{A}(x^*) \text{ with } \lambda_i^* > 0 \\ \nabla c_i(x^*)^T w \geq 0 & i \in \mathbf{I} \cap \mathbf{A}(x^*) \text{ with } \lambda_i^* = 0 \end{cases} \]

• Suppose $x^*$ is a solution to a constrained optimization problem and $\lambda^*$ satisfies KKT conditions. For all $w \in C(x^*, \lambda^*)$,

\[ w^T \nabla^2_{xx} L(x^*, \lambda^*) w \geq 0 \]
Projected Hessian

- Let $A(x^*) = [\nabla c_i(x^*)]_{i \in A(x^*)}$: $A(x^*)$ is the active set.
- It can be shown that $C(x^*, \lambda^*) = \text{Null } A(x^*)$
- $\text{Null } A(x^*)$ can be computed via QR decomposition
  
  $$A(x^*)^T = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix} = Q_1 R$$

  - $\text{span} \{ Q_2 \} = \text{Null } A(x^*)$
- Define projected Hessian $H = Q_2^T \nabla^2_{xx} L(x^*, \lambda^*) Q_2$
- The second order optimality condition is that $H$ is positive semidefinite.
Duality: An example

\[
\min_{x \geq 0} f(x) = 3x_1 + 4x_2
\]

subject to

\[
c_1 : x_1 + x_2 \geq 3
\]
\[
c_2 : 2x_1 + 0.5x_2 \geq 2
\]

- Find a lower bound for \( f(x) \) via the constraints
  - Ex: \( f(x) > 3c_1 \Rightarrow 3x_1 + 4x_2 > 3x_1 + 3x_2 \geq 9 \)
  - Ex: \( f(x) > 2c_1 + 0.5c_2 \Rightarrow 3x_1 + 4x_2 > 3x_1 + 1.25x_2 \geq 7 \)

- What is the maximum lower bound for \( f(x) \)?

\[
\max_{y \geq 0} g(y) = 3y_1 + 2y_2 \quad \text{s.t.} \quad y_1 + 2y_2 \leq 3
\]
\[
y_1 + 0.5y_2 \leq 4
\]
Dual problem

- Primal problem \( \min_x f(x) \) s.t. \( c_i(x) \geq 0 \)
  1. No equality constraints.
  2. Inequality constraints \( c_i \) are concave. \((-c_i \) are convex)

- Let \( c(x) = [c_1(x), \ldots c_m(x)]^T \). The Lagrangian is \( \mathcal{L}(x, \lambda) = f(x) - \lambda^T c(x), \quad \lambda \in \mathbb{R}^m \).

- The dual object function is \( q(\lambda) = \inf_x \mathcal{L}(x, \lambda) \)
  - If \( \mathcal{L}(\cdot, \lambda) \) is convex, \( \inf_x \mathcal{L}(x, \lambda) \) is at \( \nabla_x \mathcal{L}(x, \lambda) = 0 \)

- The dual problem is \( \max_\lambda q(x) \) s.t. \( \lambda \geq 0 \)
  - If \( \mathcal{L}(\cdot, \lambda) \) is convex, additional constraint is \( \nabla_x \mathcal{L}(x, \lambda) = 0 \)