Outline

• Lagrangian Dual

• Optimality

• Example: Dual LP
Convex Optimization

$$\min_x f(x)$$

subject to $g_i(x) \leq 0, \quad i = 1, \cdots, m$

$q_j(x) = 0, \quad j = 1, \cdots, l$

- When $f(x), g_i(x)$ are all convex and $q_j(x)$ is linear, the problem is convex optimization
- Maximizing a concave function over convex set is convex optimization problem

- Subsequently we refer this optimization problem as *primal problem*.

- A point satisfying the constraints is said to be *feasible*. 
Lagrangian and Optimality

Lagrangian: \( L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \rightarrow \mathbb{R}, \)

\[
L(x, \lambda, \nu) \overset{\text{def}}{=} f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{l} \nu_j q_j(x)
\]

- \( \lambda_i \): Lagrange multiplier associated with \( g_i(x) \leq 0 \)
- \( \nu_j \): Lagrange multiplier associated with \( q_j(x) = 0 \)

Lagrangian is central in optimization theory and algorithms.
Lagrange Dual Function

- Lagrange Dual Function: $\mathcal{G} : \mathbb{R}^m \times \mathbb{R}^l \rightarrow \mathbb{R}$,
  \[ \mathcal{G}(\lambda, \nu) \overset{\text{def}}{=} \inf_{x} L(x, \lambda, \nu) \]
  - $\lambda_i$: Lagrange multiplier associated with $g_i(x) \leq 0$
  - $\nu_j$: Lagrange multiplier associated with $q_j(x) = 0$

What properties does Lagrangian dual function have?
Note that

\[-G(\lambda, \nu) = \sup_{x} \left\{ -f(x) - \sum_{i=1}^{m} \lambda_{i}g_{i}(x) - \sum_{j=1}^{l} \nu_{j}q_{j}(x) \right\} \]

- The objective function above is linear in \((\lambda, \nu)\)
- Pointwise maximization (over \(x\)) preserves convexity
- Lagrange dual function is \textbf{concave}. 
Lagrange Dual Function

Lower Bound Property:

- If \( \mathbf{x} \) is \textit{primal feasible}, i.e., \( g_i(\mathbf{x}) \leq 0 \) and \( q_j(\mathbf{x}) = 0 \) \( \forall i, j \), and
- \( \lambda \geq 0 \), i.e., \textit{dual feasible}, then

\[
\begin{align*}
    f(\mathbf{x}) & \geq L(\mathbf{x}, \lambda, \nu) \\
    & \geq \inf_{\mathbf{x}} L(\mathbf{x}, \lambda, \nu) = \mathcal{G}(\lambda, \nu)
\end{align*}
\]
Minimizing over $\mathbf{x} \Rightarrow$

$$f^* \geq \mathcal{G}(\lambda, \nu)$$

where $f^*$ denotes the minimum value.

$\mathcal{G}(\lambda, \nu)$ is a lower bound for the optimal value $f^*$. 
Best Dual Bound

• Dual Problem:

\[
\max_{\lambda,\nu} \ G(\lambda, \nu)
\]
subject to \( \lambda \geq 0 \)

• The maximum \( G^* \) yields the best lower bound for the primal optimal \( f^* \)

• The dual problem is convex
Weak and Strong Duality

- Weak Duality: $G^* \leq f^*$
- Strong Duality: $G^* = f^*$

- If duality gap is zero at a primal dual feasible point $(x^*, \lambda^*, \nu^*)$, primal and dual optimality is attained.
- For a convex problem, strong duality usually holds (assuming existence of strict feasible point wrt inequality constraints)
- For nonconvex problem, strong duality may not hold
Dual LP: an example

• Standard LP

\[
\min_x c^T x \\
\text{subject to } Ax = b \\
x \geq 0
\]
• Lagrangian Function:

\[ L(x, \lambda, \nu) = c^T x - \lambda^T x + \nu^T (A x - b) \]
Dual LP

• Dual Function:

$$G(\lambda, \nu) = \inf_x (c^T x - \lambda^T x + \nu^T (Ax - b))$$

$$= \inf_x ( (c - \lambda + A^T \nu)^T x - b^T \nu )$$

$$= \begin{cases} 
-\nu^T b & \text{if } A^T \nu - \lambda + c = 0 \\
-\infty & \text{otherwise} 
\end{cases}$$
• Dual LP

\[
\max_{\nu} \quad -b^T \nu \\
\text{subject to} \quad A^T \nu - \lambda + c = 0 \quad \text{(DLP)} \\
\lambda \geq 0
\]
Karush-Kuhn-Tucker (KKT) Conditions

Assume $f, q, g$ are differentiable.

$$\min_x f(x)$$

subject to $q_i(x) = 0, \ i = 1, \cdots, l$ \hspace{1cm} (NLP)

$g_i(x) \leq 0, \ i = 1, \cdots, m$

Assume strict duality holds at feasible $(x^*, \lambda^*, \nu^*)$

$$f(x^*) = G(\lambda^*, \nu^*) = \inf_x L(x, \lambda^*, \nu^*)$$

$$\leq f(x^*) + \sum_{i=1}^{m} \lambda^*_i g_i(x^*) + \sum_{j=1}^{l} \nu^*_j q_j(x^*)$$

$$\leq f(x^*) \hspace{1cm} \text{(due to feasibility)}$$

$\Rightarrow$ inequalities must be equalities
Optimality Conditions

KKT Conditions: Let \( f, q, g \) be continuously differentiable.

- **Primal Feasibility:**
  \[ q_i(x^*) = 0, \quad i = 1, \ldots, l, \quad g_i(x^*) \leq 0, \quad i = 1, \ldots, m \]

- **Dual Feasibility:** \( \lambda^* \geq 0 \)

- **Gradient with respect to** \( x \), \( \nabla_x L(\cdot) \), **equal zero:**
  \[ \nabla_x L(x^*, \lambda^*, \nu^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{l} \nu_j^* \nabla q_j(x^*) = 0 \quad (1) \]

- **Complementarity:** \( \lambda_i^* g_i(x^*) = 0, \quad i = 1, \ldots, m \)
Optimality Conditions

- For a convex problem with differentiable objective and constraints satisfying a technical condition, KKT is necessary and sufficient.

- For a nonconvex problem with differentiable objective and constraints, KKT can only be a necessary condition.