Overview

- Duality
  - Lagrange dual
  - KKT condition
Standard form for constrained nonlinear optimization

\[
\begin{align*}
\min_{X} & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
& \quad h_n(X) = 0 \quad (n = 1, 2, \ldots, N)
\end{align*}
\]

We do not write equality constraint \( h(X) = 0 \) as two inequality constraints \( h(X) \geq 0 \) and \( h(X) \leq 0 \) in this lecture.

Equality and inequality constraints are handled differently in duality theory.
Define the Lagrangian

\[
\text{min}_{X} \quad f(X) \\
\text{S.T.} \quad g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
\quad h_n(X) = 0 \quad (n = 1, 2, \ldots, N)
\]

\[
L(X, U, V) = f(X) + \sum_{m=1}^{M} u_m g_m(X) + \sum_{n=1}^{N} v_n h_n(X)
\]

\(L(X, U, V)\) is a nonlinear function of \(X\), but it is **linearly** dependent of \(U\) and \(V\)
Define **Lagrange dual function**

\[
d(U,V) = \inf_X L(X,U,V) = \inf_X \left[ f(X) + \sum_{m=1}^{M} u_m g_m(X) + \sum_{n=1}^{N} v_n h_n(X) \right]
\]

At any given \( X \), \( L(X,U,V) \) is a linear function of \( U \) and \( V \)

\( d(U,V) \) is the minimum of an infinite number of linear functions
For any constrained nonlinear optimization, the Lagrange dual function $d(U,V)$ is concave.
Lower Bound Property

\[
\begin{array}{ll}
\min_{X} & f(X) \\
\text{S.T.} & g_m(X) \leq 0 \quad (m = 1, 2, \cdots, M) \\
& h_n(X) = 0 \quad (n = 1, 2, \cdots, N) \\
\end{array}
\]

\[
d(U, V) = \inf_{X} \left[ f(X) + \sum_{m=1}^{M} u_m g_m(X) + \sum_{n=1}^{N} v_n h_n(X) \right]
\]

- If \( X^* \) is the optimal solution and \( U \geq 0 \), then

\[
\begin{align*}
g_m(X^*) & \leq 0 \quad (m = 1, 2, \cdots, M) \\
h_n(X^*) & = 0 \quad (n = 1, 2, \cdots, N) \\
\end{align*}
\]

\[
d(U, V) \leq f(X^*) + \sum_{m=1}^{M} u_m g_m(X^*) + \sum_{n=1}^{N} v_n h_n(X^*)
\]

\[
= f(X^*) + \sum_{m=1}^{M} u_m g_m(X^*)
\]

\[
\leq f(X^*) \quad \text{d(U,V) is the lower bound of f(X*)}
\]
Linear Programming Example

\[
\min_{X} \quad C^T X \\
\text{S.T.} \quad AX = B \quad X \leq 0
\]

\[
L(X,U,V) = C^T X + U^T X + V^T \cdot (AX - B) \\
= (C^T + U^T + V^T A) \cdot X - V^T B
\]

\[
d(U,V) = \inf_{X} L(X,U,V) = \begin{cases} 
-V^T B & \text{if } (C^T + U^T + V^T A = 0) \\
-\infty & \text{(Otherwise)}
\end{cases}
\]

Concave function

\[
C^T X^* \geq -V^T B \quad \left( C^T + U^T + V^T A = 0 \quad U \geq 0 \right)
\]
Lagrange Dual Problem

- **Lagrange dual problem** is defined as

\[
\begin{align*}
\min_x & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
& \quad h_n(X) = 0 \quad (n = 1, 2, \ldots, N)
\end{align*}
\]

Primal problem

\[
\begin{align*}
\max_{U,V} & \quad d(U,V) \\
\text{S.T.} & \quad U \geq 0
\end{align*}
\]

Dual problem

- **Linear programming example**

\[
\begin{align*}
\min_x & \quad C^T X \\
\text{S.T.} & \quad AX = B \\
& \quad X \leq 0
\end{align*}
\]

\[
\begin{align*}
\max_{U,V} & \quad -V^T B \\
\text{S.T.} & \quad C^T + U^T + V^T A = 0 \\
& \quad U \geq 0
\end{align*}
\]
Weak Duality

\[
\begin{align*}
\min_{X} & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1,2,\ldots,M) \\
& \quad h_n(X) = 0 \quad (n = 1,2,\ldots,N)
\end{align*}
\]

Primal problem

- Weak duality
  - \(X^*\) is primal optimum
  - \(U^*\) and \(V^*\) are dual optimum
  - \(f(X^*) \geq d(U^*,V^*)\) (Lagrange dual function is the lower bound)

- Weak duality holds for any optimization problem (either convex or non-convex)

\[
\begin{align*}
\max_{U,V} & \quad d(U,V) \\
\text{S.T.} & \quad U \geq 0
\end{align*}
\]

Dual problem
Strong Duality

\[
\begin{align*}
\min_{X} & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
& \quad h_n(X) = 0 \quad (n = 1, 2, \ldots, N)
\end{align*}
\]
Primal problem

\[
\begin{align*}
\max_{U, V} & \quad d(U, V) \\
\text{S.T.} & \quad U \geq 0
\end{align*}
\]
Dual problem

- **Strong duality**
  - $X^*$ is primal optimum
  - $U^*$ and $V^*$ are dual optimum
  - $f(X^*) = d(U^*, V^*)$ (duality gap is zero)

- **Strong duality does not hold in general, but it usually holds for convex problems**
  - Conditions that guarantee strong duality in convex problems are referred to as constraint qualifications
Slater’s Constraint Qualification

- **Strong duality holds for convex optimization**

  \[
  \min_X f(X) \quad \text{S.T.} \quad g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
  AX = B
  \]

  if it is strictly feasible, i.e.,

  \[
  g_m(X) < 0 \quad (m = 1, 2, \ldots, M) \\
  AX = B
  \]

- **Sufficient but not necessary condition**

  - Many other constraint qualifications exist
Quadratic Programming Example

Primal problem

\[
\begin{align*}
\min_x & \quad X^T AX + 2B^T X \\
\text{S.T.} & \quad X^T X \leq 1
\end{align*}
\]

Dual problem

\[
\begin{align*}
\max_{t,u} & \quad -t - u \\
\text{S.T.} & \quad \begin{bmatrix} A + uI & B \\ B^T & t \end{bmatrix} \succeq 0 \\
& \quad u \geq 0
\end{align*}
\]

- Primal problem is not convex, if A is not positive semidefinite
- Dual problem is convex semidefinite programming
- Strong duality holds even if primal problem is not convex
  - Dual problem can be solved both efficiently and robustly due to convexity
Complementary Slackness

\[ \min_{X} f(X) \]

\[ \text{S.T. } g_m(X) \leq 0 \quad (m = 1,2,\ldots,M) \]

\[ h_n(X) = 0 \quad (n = 1,2,\ldots,N) \]

Primal problem

\[ \max_{U,V} d(U,V) \]

\[ \text{S.T. } U \geq 0 \]

Dual problem

Assume that strong duality holds, \( X^* \) is primal optimum, and \( U^* \) and \( V^* \) are dual optimum

\[
f(X^*) = d(U^*,V^*) = \inf_{X} \left( f(X) + \sum_{m=1}^{M} u^*_m g_m(X) + \sum_{n=1}^{N} v^*_n h_n(X) \right) \]

\[
\leq f(X^*) + \sum_{m=1}^{M} u^*_m g_m(X^*) + \sum_{n=1}^{N} v^*_n h_n(X^*)
\]

\[
= f(X^*) + \sum_{m=1}^{M} u^*_m g_m(X^*)
\]

\[
\leq f(X^*)
\]
Complementary Slackness

\[
\min_{X} f(X) \\
\text{S.T. } g_m(X) \leq 0 \quad (m = 1, 2, \ldots, M) \\
\quad h_n(X) = 0 \quad (n = 1, 2, \ldots, N)
\]

Primal problem

\[
f(X^*) \leq f(X^*) + \sum_{m=1}^{M} u^*_m g_m(X^*) \leq f(X^*)
\]

\[
\sum_{m=1}^{M} u^*_m g_m(X^*) = 0 \quad u^*_m g_m(X^*) \leq 0
\]

\[u^*_m g_m(X^*) = 0\]

- \(u^*_m > 0 \rightarrow g_m(X^*) = 0\) (active constraint)
- \(g_m(X^*) < 0 \rightarrow u^*_m = 0\) (inactive constraint)
Karush-Kuhn-Tucker (KKT) Condition

\[
\begin{align*}
\min_{X} & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1,2,\ldots, M) \\
& \quad h_n(X) = 0 \quad (n = 1,2,\ldots, N)
\end{align*}
\]

Primal problem

\[
\begin{align*}
\max_{U,V} & \quad d(U, V) \\
\text{S.T.} & \quad U \geq 0
\end{align*}
\]

Dual problem

If strong duality holds and \( X^*, U^* \) and \( V^* \) are optimal, then

\[
\begin{align*}
g_m(X^*) & \leq 0 \quad (m = 1,2,\ldots, M) \\
h_n(X^*) & = 0 \quad (n = 1,2,\ldots, N)
\end{align*}
\]

Primal constraints

\[
U^* \geq 0
\]

Dual constraints

\[
u_m^* g_m(X^*) = 0 \quad (m = 1,2,\ldots, M)
\]

Complementary slackness

\[
\nabla f(X^*) + \sum_{m=1}^{M} u_m^* \cdot \nabla g_m(X^*) + \sum_{n=1}^{N} v_n^* \cdot \nabla h_n(X^*) = 0
\]

\( X^* \) minimizes \( L(X,U^*,V^*) \)
KKT Condition for Convex Problem

\[
\begin{align*}
\min_{X} & \quad f(X) \\
\text{S.T.} & \quad g_m(X) \leq 0 \quad (m = 1,2,\ldots, M) \\
& \quad h_n(X) = 0 \quad (n = 1,2,\ldots, N)
\end{align*}
\]

Primal problem

\[
\begin{align*}
\max_{U,V} & \quad d(U,V) \\
\text{S.T.} & \quad U \geq 0
\end{align*}
\]

Dual problem

- Given a convex problem with strong duality, \( X^*, U^* \) and \( V^* \) are optimal if and only if they satisfy the KKT condition.

- Many convex programming algorithms are derived from KKT.

Summary

- Duality
  - Lagrange dual
  - KKT condition