18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213

Overview

- Duality
  - Lagrange dual
  - KKT condition

Constrained Nonlinear Optimization

- Standard form for constrained nonlinear optimization

$$
\begin{aligned}
\min_{X} \quad & f(X) \\
S.T. \quad & g_m(X) \le 0 \quad (m = 1,2,\cdots,M) \\
& h_n(X) = 0 \quad (n = 1,2,\cdots,N)
\end{aligned}
$$

- We do not write equality constraint h(X) = 0 as two inequality constraints h(X) ≥ 0 and h(X) ≤ 0 in this lecture
  - Equality and inequality constraints are handled differently in duality theory

Lagrangian

$$\begin{aligned} \min_X \quad & f(X) \\ S.T. \quad & g_m(X) \le 0 \quad (m = 1, 2, \cdots, M) \\ & h_n(X) = 0 \quad (n = 1, 2, \cdots, N) \end{aligned}$$

- Define the Lagrangian

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

Lagrange multipliers

  - L(X,U,V) is a nonlinear function of X, but it is linearly dependent of U and V

Lagrange Dual Function

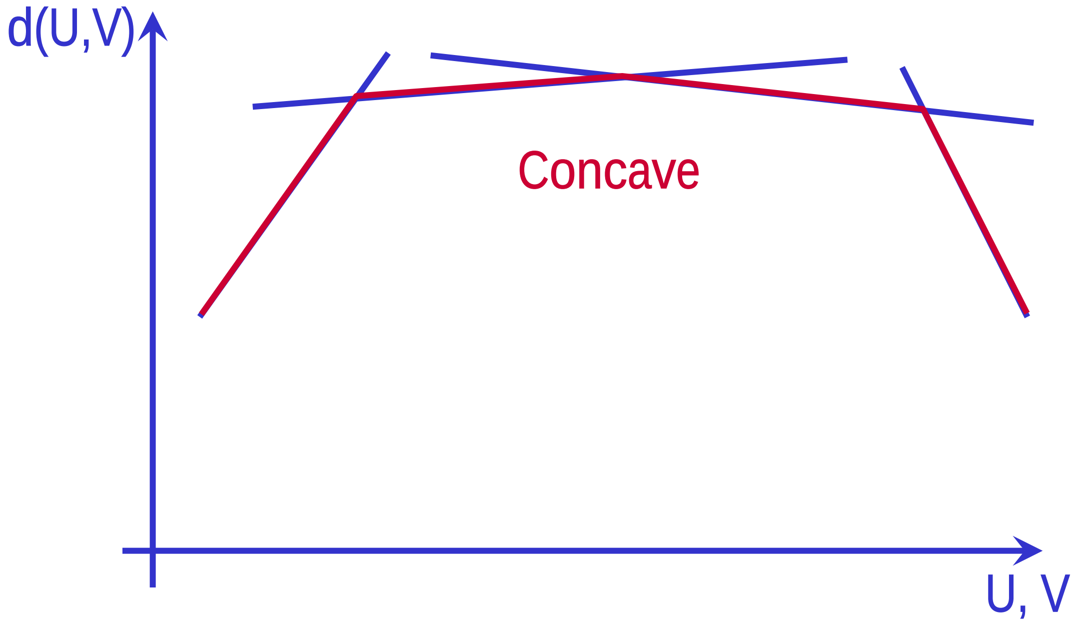
- Define Lagrange dual function

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

- 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

Lagrange Dual Function

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



- For any constrained nonlinear optimization, the Lagrange dual function d(U,V) is concave

Lower Bound Property

$$\begin{aligned} \min_X \quad & f(X) \\ \text{S.T.} \quad & g_m(X) \le 0 \quad (m = 1,2,\cdots,M) \\ & h_n(X) = 0 \quad (n = 1,2,\cdots,N) \end{aligned}$$

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

- If X* is the optimal solution and U ≥ 0, then

$$\begin{aligned} & g_m(X^*) \le 0 \quad (m = 1,2,\cdots,M) \\ & h_n(X^*) = 0 \quad (n = 1,2,\cdots,N) \end{aligned}$$

$$\begin{aligned} d(U,V) &\le f(X^*) + \sum_{m=1}^{M} u_m g_m(X^*) + \sum_{n=1}^{N} v_n h_n(X^*) \\ &= f(X^*) + \sum_{m=1}^{M} u_m g_m(X^*) \\ &\le f(X^*) \end{aligned}$$

d(U,V) is the lower bound of f(X*)

Linear Programming Example

$$\begin{aligned} \min_{X} \quad & C^T X \\ \text{S.T.} \quad & AX = B \\ & X \le 0 \end{aligned}$$

$$\begin{aligned} L(X,U,V) &= C^T X + U^T X + V^T \cdot (AX - B) \\ &= \left(C^T + U^T + V^T A\right) \cdot X - V^T B \end{aligned}$$

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

Concave function

$$C^T X^* \ge -V^T B \quad \left(C^T + U^T + V^T A = 0 \quad U \ge 0\right)$$

Lagrange Dual Problem

- Lagrange dual problem is defined as

$$\begin{aligned} \min_X \quad & f(X) \\ \text{S.T.} \quad & g_m(X) \le 0 \quad (m = 1, 2, \cdots, M) \\ & h_n(X) = 0 \quad (n = 1, 2, \cdots, N) \end{aligned}$$

Primal problem

$$\Rightarrow$$

$$\begin{aligned} \max_{U,V} \quad & d(U, V) \\ \text{S.T.} \quad & U \ge 0 \end{aligned}$$

Dual problem

- Linear programming example

$$\begin{aligned} \min_X \quad & C^T X \\ \text{S.T.} \quad & AX = B \\ & X \le 0 \end{aligned}$$

$$\Rightarrow$$

$$\begin{aligned} \max_{U,V} \quad & -V^T B \\ \text{S.T.} \quad & C^T + U^T + V^T A = 0 \\ & U \ge 0 \end{aligned}$$

Weak Duality

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad g_m(X) \le 0 \quad (m = 1,2,\cdots,M)$$
$$h_n(X) = 0 \quad (n = 1,2,\cdots,N)$$

Primal problem

$$\max_{U,V} \quad d(U,V)$$
$$\text{S.T.} \quad U \ge 0$$

Dual problem

- Weak duality
  - X* is primal optimum
  - U* and V* are dual optimum
  - f(X*) ≥ d(U*,V*) (Lagrange dual function is the lower bound)
- Weak duality holds for any optimization problem (either convex or non-convex)

Strong Duality

$$\begin{aligned} \min_{X} \quad & f(X) \\ \text{S.T.} \quad & g_m(X) \le 0 \quad (m = 1,2,\cdots,M) \\ & h_n(X) = 0 \quad (n = 1,2,\cdots,N) \end{aligned}$$

Primal problem

$\Rightarrow$

$$\begin{aligned} \max_{U,V} \quad & d(U,V) \\ \text{S.T.} \quad & U \ge 0 \end{aligned}$$

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

$$\begin{aligned} &\min_X \quad f(X) \\ &\text{S.T.} \quad g_m(X) \le 0 \quad (m = 1,2,\cdots,M) \\ &\qquad\quad AX = B \end{aligned}$$

Equality constraints must be linear

  - if it is strictly feasible, i.e.,

$$\begin{aligned} &g_m(X) < 0 \quad (m = 1,2,\cdots,M) \\ &AX = B \end{aligned}$$

- Sufficient but not necessary condition
  - Many other constraint qualifications exist

Quadratic Programming Example

$$\min_{X} \quad X^T A X + 2B^T X$$
$$\text{S.T.} \quad X^T X \leq 1$$

Primal problem

$$\max_{t,u} \quad -t-u$$
$$\text{S.T.} \quad \begin{bmatrix} A+uI & B \\ B^T & t \end{bmatrix} \succ= 0$$
$$u \geq 0$$

Dual problem

- 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

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad g_m(X) \le 0 \quad (m = 1,2,\cdots,M)$$
$$h_n(X) = 0 \quad (n = 1,2,\cdots,N)$$

Primal problem

$$\max_{U,V} \quad d(U,V)$$
$$\text{S.T.} \quad U \ge 0$$

Dual problem

- Assume that strong duality holds, X* is primal optimum, and U* and V* are dual optimum

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

Complementary Slackness

$$\min_X \quad f(X)$$
$$\text{S.T.} \quad g_m(X) \le 0 \quad (m = 1,2,\cdots,M)$$
$$h_n(X) = 0 \quad (n = 1,2,\cdots,N)$$

Primal problem

$\Rightarrow$

$$\max_{U,V} \quad d(U,V)$$
$$\text{S.T.} \quad U \ge 0$$

Dual problem

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

$$\sum_{m=1}^{M} u_m^* g_m(X^*) = 0 \quad u_m^* g_m(X^*) \le 0$$

$$u_m^* g_m(X^*) = 0$$

- $u_m^* > 0 \rightarrow g_m(X^*) = 0$ (active constraint)
- $g_m(X^*) < 0 \rightarrow u_m^* = 0$ (inactive constraint)

Karush-Kuhn-Tucker (KKT) Condition

$$\min_{X} \quad f(X)$$
$$\text{S.T.} \quad g_m(X) \le 0 \quad (m = 1,2,\cdots,M)$$
$$h_n(X) = 0 \quad (n = 1,2,\cdots,N)$$

Primal problem

$$\Rightarrow$$

$$\max_{U,V} \quad d(U,V)$$
$$\text{S.T.} \quad U \ge 0$$

Dual problem

- If strong duality holds and X*, U* and V* are optimal, then

$$g_m(X^*) \le 0 \quad (m = 1,2,\cdots,M)$$
$$h_n(X^*) = 0 \quad (n = 1,2,\cdots,N)$$

Primal constraints

$$U^* \ge 0$$

Dual constraints

$$u_m^* g_m(X^*) = 0 \quad (m = 1,2,\cdots,M)$$

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

$$\begin{aligned}\min_{X} \quad & f(X) \\ \text{S.T.} \quad & g_m(X) \le 0 \quad (m = 1,2,\cdots,M) \\ & h_n(X) = 0 \quad (n = 1,2,\cdots,N)\end{aligned}$$

Primal problem

$$\Rightarrow$$

$$\begin{aligned}\max_{U,V} \quad & d(U,V) \\ \text{S.T.} \quad & U \ge 0\end{aligned}$$

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

Boyd and Vandenberghe, “Convex Optimization,” Cambridge University Press, 2004

Summary

- Duality
  - Lagrange dual
  - KKT condition