ECE236B Discussion 6

Xin Jiang

ECE236B Convex Optimization (Winter 2021)

Feburary 12, 2021

Problem setup

Standard form problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- domain of primal problem $\mathcal{D} = \cap_{i=0}^{m} \operatorname{dom} f_i \bigcap \cap_{i=1}^{p} \operatorname{dom} h_i$
- variable $x \in \mathbf{R}^n$, optimal value $p^\star$

Lagrangian $L\colon \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\operatorname{dom} L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

Lagrange dual

Lagrange dual function: the infimum of $L(x, \lambda, \nu)$

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

$$= \begin{cases} \cdots & (\lambda, \nu) \in \operatorname{dom} g \\ -\infty & \text{otherwise} \end{cases}$$

Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array} \qquad \begin{array}{ll} \text{maximize} & \cdots \\ \text{subject to} & \lambda \succeq 0 \\ & (\lambda, \nu) \in \operatorname{dom} g \end{array}$$

- a convex optimization problem; optimal value $d^\star$
- (λ, ν) is *dual feasible* if $\lambda \succeq 0$ and $(\lambda, \nu) \in \operatorname{dom} g$

Two-way partition example

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1 \end{array} \quad \text{(primal)}$$

$$\begin{array}{ll} \text{maximize} & g(\nu) = -\mathbf{1}^T \nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0 \end{array} \quad \text{(dual)}$$

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & \sum x_i^2 = n \end{array} \quad \text{(relax)}$$

$$\begin{array}{ll} \text{minimize} & \mathrm{tr}(WX) \\ \text{subject to} & X_{ii} = 1,\ X \succeq 0 \end{array} \quad \text{(dual 2)}$$

$$\begin{array}{ll} \text{minimize} & \mathrm{tr}(WX) \\ \text{subject to} & X_{ii} = 1 \\ & X \succeq 0,\ \mathrm{rank}\, X = 1 \end{array} \quad \text{(primal SDP)}$$

Two-way partition example cont'd

- (primal) is not convex when $W \notin \mathbf{S}_+^n$
- (relax) is a relaxation of (primal) since we enlarge the feasible region
- (dual) is the dual of (primal), and (dual 2) is the dual of (dual)
- (primal SDP) is an equivalent reformulation by defining $X = xx^T$
- any positive semidefinite rank-one matrix has the form $X = xx^T$, the converse is also true
- $p^\star$, $d^\star$, $q^\star$ are the optimal values for (primal), (dual), and (relax), respectively
- $\tilde{\nu} = -\lambda_{\min}(W) \cdot \mathbf{1}$ is feasible for (dual), with $g(\tilde{\nu}) = n\lambda_{\min}(W)$

$$p^\star \geq d^\star \geq g(\tilde{\nu}) = q^\star$$

- (dual 2) is a relaxation of (primal) or (primal SDP)

Strong duality and Slater's condition

Weak duality $d^\star \leq p^\star$

- weak duality always holds, even for nonconvex problems
- weak duality can be meaningless when $g(\lambda, \nu) = -\infty$

Strong duality $p^\star = d^\star$

- convex problems $\not\Leftrightarrow$ strong duality
- convex problems + constraint qualification $\Longrightarrow$ strong duality
- Slater's theorem: for a convex problem, if there exists a strictly feasible x, then
  1. strong duality holds, *i.e.*, $p^\star = d^\star$;
  2. furthermore, if $p^\star > -\infty$, then the dual optimum is attained, *i.e.*, there exists $(\lambda^\star, \nu^\star)$ with $g(\lambda^\star, \nu^\star) = d^\star = p^\star$

Example: convex problem vs. strong duality

A convex problem without strong duality (T5.21)

$$\begin{array}{ll} \text{minimize} & e^{-x} \\ \text{subject to} & f_1(x,y) \le 0 \end{array}$$

where $f_1(x,y) = \begin{cases} x^2/y, & y > 0 \\ \infty, & \text{otherwise} \end{cases}$

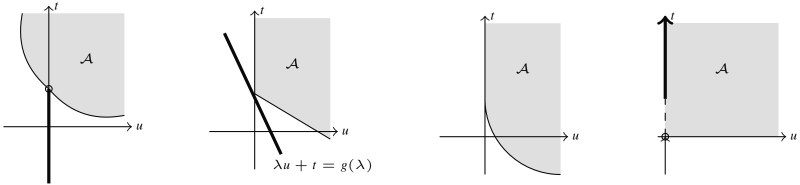
$$\begin{array}{ll} \text{maximize} & 0 \\ \text{subject to} & \lambda \ge 0 \end{array}$$

A nonconvex problem with strong duality

$$\begin{array}{ll} \text{minimize} & x^T A x + 2b^T x \\ \text{subject to} & x^T x \le 1 \end{array}$$

$$\begin{array}{ll} \text{maximize} & -t - \lambda \\ \text{subject to} & \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \succeq 0 \end{array}$$

Geometric interpretation



$$\mathcal{A} = \{(u,t) \in \mathbf{R}^2 \mid f_1(x) \le u, f_0(x) \le t \text{ for some } x \in \mathcal{D}\}$$

- $\mathcal{A}$ is convex for convex problems; a supporting hyperplane at $(0, p^\star)$
- Slater's condition: if there exists a pair of $(\tilde{u}, \tilde{t}) \in \mathcal{A}$ with $\tilde{u} < 0$, then supporting hyperplanes at $(0, p^\star)$ must be non-vertical

Slater's condition	$\surd$	$\times$	$\times$	$\times$
strong duality	$\surd$	$\surd$	$\surd$	$\times$
dual optimal attained	$\surd$	$\surd$	$\times$	$\surd$

Optimality conditions

we say the pair (x, λ, ν) satisfies the optimality conditions (OC) if

1. primal constraints: $f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p$
2. dual constraints: $\lambda \succeq 0$ (dual "implicit" constraint: $(\lambda, \nu) \in \mathbf{dom}\, g$)
3. complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$
4. x is a minimizer of $L(\cdot, \lambda, \nu)$

4'. gradient of Lagrangian with respect to x vanishes

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

primal and dual optimal, strong duality $\iff$ optimality conditions (with #4)

if the problem is convex, differentiable, and satisfies Slater's condition

primal and dual optimal $\iff$ optimality conditions with #4' (or KKT)

conditions 1, 2, 3, and 4' are known as *Karush–Kuhn–Tucker* (KKT) conditions

Optimality conditions cont'd

For nonconvex problems

- the pair of points (x, λ, ν) satisfy optimality conditions if and only if

  1. x is primal optimal 2. (λ, ν) is dual optimal 3. strong duality holds

- the pair of points (x, λ, ν) satisfy optimality conditions if and only if

  1. x is primal feasible 2. (λ, ν) is dual feasible 3. $f_0(x) = g(\lambda, \nu)$

For convex, differentiable problems satisfying Slater's condition
(in this case, #4 and #4' are equivalent)

- the pair of points (x, λ, ν) satisfy KKT condition if and only if

  x is primal optimal, and (λ, ν) is dual optimal

- if x is primal optimal, there exists a dual feasible point s.t. $f_0(x) = g(\lambda, \nu)$

Example: optimality conditions

KKT (with #4') is not sufficient for optimality in a nonconvex problem (T5.29)

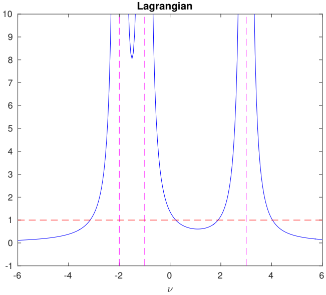
$$\begin{array}{ll} \text{minimize} & -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) \\ \text{subject to} & x_1^2 + x_2^2 + x_3^2 = 1 \end{array}$$

L is a quadratic function in x:

$$L(x, \nu) = (-3 + \nu)x_1^2 + (1 + \nu)x_2^2 + (2 + \nu)x_3^2 + 2(x_1 + x_2 + x_3) - \nu$$

KKT condition:

$$\frac{1}{(-3 + \nu)^2} + \frac{1}{(1 + \nu)^2} + \frac{1}{(2 + \nu)^2} = 1$$



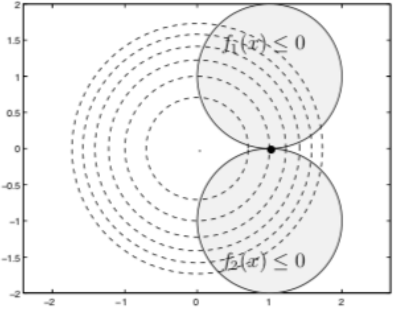
$\tilde{x}$ is a minimizer of $L(x, \tilde{\nu})$ if $\nabla L(\tilde{x}, \tilde{\nu}) = 0$ and $\nabla^2 L(\tilde{x}, \tilde{\nu}) \succeq 0$; hence, all the coefficients before the quadratic terms must be nonnegative, and thus the optimal $\tilde{\nu}$ must be greater than 3

Example: Optimality conditions

KKT is not necessary for optimality when Slater's condition fails (T5.26)

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \\ & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1 \end{array}$$

- Slater's condition fails
- strong duality holds
- dual optimum is not attained
- KKT system is not solvable



Semidefinite program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_{i=1}^n x_i F_i \preceq G \end{array} \qquad \begin{array}{ll} \text{maximize} & -\operatorname{tr}(GZ) \\ \text{subject to} & \operatorname{tr}(F_i Z) + c_i = 0 \\ & Z \succeq 0 \end{array}$$

strong duality does not always hold for SDPs

- SDP with positive duality gap
- SDP with zero duality gap, but no attainment

Strong duality holds if primal or dual SDP is strictly feasible.

Example: SDP with positive duality gap

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & \begin{bmatrix} 0 & x_1 \\ x_1 & x_2 \end{bmatrix} \succeq 0, \quad x_1 \geq -1 \end{array}$$

primal optimal is $p^\star = 0$; Slater's condition is not satisfied

$$\begin{aligned} L(x, Z, z) &= x_1 - \operatorname{tr}\left(Z \begin{bmatrix} 0 & x_1 \\ x_1 & x_2 \end{bmatrix} \right) - z(x_1 + 1) \quad (Z \succeq 0, z \geq 0) \\ &= (1 - z - 2Z_{12})x_1 - x_2 Z_{22} - z \end{aligned}$$

the dual problem is

$$\begin{array}{ll} \text{maximize} & -z \\ \text{subject to} & \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{bmatrix} \succeq 0 \\ & z \geq 0 \\ & 1 - z - 2Z_{12} = 0 \\ & Z_{22} = 0 \end{array}$$

dual optimal is $d^\star = -1$; hence strong duality does not hold

Example: SDP with zero duality gap

$$\begin{array}{ll} \text{minimize} & X_{11} \\ \text{subject to} & 2X_{12} = 1, \quad X \succeq 0, \end{array}$$

where the variable is $X \in \mathbf{S}^2$; note that $X_{11} > 0$ for all feasible X, and the series

$$X^{(k)} = \begin{bmatrix} 1/k & 1/2 \\ 1/2 & k \end{bmatrix}$$

is feasible and satisfies $X_{11}^{(k)} \to 0$; so $p^\star = 0$ but optimal solution is not attained

the dual SDP is

$$\begin{array}{ll} \text{maximize} & y \\ \text{subject to} & y \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

the only feasible point is $y = 0$; so the optimal value is $d^\star = 0$ and is attained

Conclusion: strong duality holds, but primal optimal is not attained