Duality theory

G. Ferrari Trecate

Dipartimento di Ingegneria Industriale e dell'Informazione
Università degli Studi di Pavia

Industrial Automation

Motivation

For unconstrained convex optimization problems, minima can be computed solving algebraic equations

$$\nabla f(x^*) = 0$$

GOAL: provide algebraic conditions also for constrained problems

Outline:

- Duality theory
- Main results for convex programming problems
- Duality theory for LP

Preliminaries

Reference optimization problem

$$J^* = \min f(x) \qquad (P_{min})$$
$$g(x) \leq 0$$
$$h(x) = 0$$

where $g(x) \in \mathbb{R}^m$, $h(x) \in \mathbb{R}^p$ and $J^* \in [-\infty, +\infty]$

- Feasible region $X = \{x \in \mathbb{R}^n : g(x) \leq 0 \text{ and } h(x) = 0\}$

Every $x_0 \in X$ gives an upper bound to J^* because $f(x_0) \geq J^*$

How to find a lower bound? Duality theory!

Lagrangian function

If $x \in X$, for all $\lambda \in \mathbb{R}^n$, $\lambda \geq 0$ and $\mu \in \mathbb{R}^p$ one has

$$L(x,\lambda,\mu) \triangleq f(x) + \underbrace{\lambda' g(x)}_{\leq 0} + \underbrace{\mu' h(x)}_{=0} \leq f(x)$$

Then, $\forall \mu$, $\forall \lambda \geq 0$

$$\min_{x} L(x,\lambda,\mu) \leq \min_{x \in X} L(x,\lambda,\mu) \leq \min_{x \in X} f(x) = J^*$$

$$\downarrow$$
$$X \subseteq \mathbb{R}^n$$

Take the maximal lower bound and obtain

$$\max_{\lambda \geq 0, \mu} \left[\min_{x} L(x,\lambda,\mu) \right] \leq J^*$$

Jargon

- λ, μ : Lagrange multipliers
- $L(x, \lambda, \mu)$: Lagrangian function
- $\ell(\lambda, \mu) = \min_x L(x, \lambda, \mu)$: dual cost

$$D^* = \max_{\lambda \geq 0, \mu} \ell(\lambda, \mu) \qquad (D\text{-pmin})$$

Rmk

1) $\ell(\lambda, \mu)$ produced by an unconstrained optimization problem
2) $\ell(\lambda, \mu)$ is concave, irrespectively of f, g, h

   $\hookrightarrow$ (D-pmin) is a convex programming problem
3) (P_{min}) is called primal problem and its optimizers are primal optimizers

Def. (λ, μ) is feasible for (D-prim) if $\lambda \geq 0$ and $\ell(\lambda, \mu) > -\infty$

Thm. The dual problem of the dual problem is the primal problem

Def. $D^* \leq Z^*$ is the weak duality relation
$D^* = Z^*$ is the strong duality relation

Problem: Weak duality holds by construction. When strong duality also holds?

Strong duality for convex programming

Def. (P_{min}) is a convex programming problem if f and g are convex and h is affine

Def. In a convex programming problem, constraints are qualified if (P_{min}) is feasible and

$$\exists x_0 \in X : g_J(x_0) < 0 \text{ if } g_J \text{ is not affine} \qquad (*)$$

Rmk.
- g_J all affine $\Rightarrow$ qualification and feasibility coincide
- $(*)$ are called Slater's conditions

KT points

Def. A point (λ^*, μ^*) is a Kuhn-Tucker (KT) point if

$$\lambda^* \geq 0 \quad \text{and} \quad \ell(\lambda^*, \mu^*) = \mathfrak{Z}^* \qquad (KTC)$$

Rmk. • Since $D^* \leq \mathfrak{Z}^*$, the existence of a KT point guarantees that $D^* = \mathfrak{Z}^*$ (strong duality)

• If (P_{min}) is unbounded, $\mathfrak{Z}^* = -\infty$ and then $D^* = -\infty$, i.e. the dual problem is infeasible

Thm (strong duality for convex programming). If (P_{min}) is convex and constraints are qualified, there is a KT point

Relations between primal and dual optimizers

Thm. If x^* is a primal optimizer and (λ^*, μ^*) is a KT point, then

$$\text{(a)} \quad x^* \text{ is a minimizer of } L(x, \lambda^*, \mu^*)$$

$$\text{(b)} \quad \lambda_i^* g_i(x^*) = 0 \qquad i = 1, \dots, m$$

Proof of point (a). (λ^*, μ^*) is a KT point $\Rightarrow$ $\underbrace{\min_x L(x, \lambda^*, \mu^*)}_{\ell(\lambda^*, \mu^*)} = \mathfrak{J}^*$

Then, $L(x, \lambda^*, \mu^*) \geq \overbrace{f(x^*)}^{\mathfrak{J}^*}$, $\forall x \in X$. Moreover

$$f(x^*) \geq L(x^*, \lambda^*, \mu^*) \longrightarrow \text{by construction of } L$$

and then, $L(x, \lambda^*, \mu^*) \geq L(x^*, \lambda^*, \mu^*)$, $\forall x \in X$, that is point (a)

Proof of point (b). From point (a) and the definition of a KT point

$$\underbrace{L(x^*,\lambda^*,\mu^*)}_{\ell(\lambda^*,\mu^*)} = \underbrace{f(x^*)}_{\zeta^*}$$

that is

$$f(x^*) + \lambda^{*T} g(x^*) + \underbrace{\mu^{*T} h(x^*)}_{=0} = f(x^*)$$

Then,

$$\lambda^{*T} g(x^*) = 0 \Rightarrow \sum_{i=1}^{m} \lambda_i^* g_i(x^*) = 0 \Rightarrow \lambda_i^* g_i(x^*) = 0 \quad i = 1,..,m$$

Remarks

1) The theorem says

$$\bar{x} = \arg\min_{x \in X} f(x) \Rightarrow \bar{x} = \arg\min L(x, \mu^*, \lambda^*)$$

but NOT the viceversa. There might be minimizers of $L(x, \mu^*, \lambda^*)$ that are NOT minimizers of $f(x)$. Below, we will discuss when the opposite implication holds

2) The theorem holds even if (P_{min}) is not convex

3) Relations (b) are called complementarity slackness conditions

$$\text{If } \lambda_i^* g_i(x^*) = 0, \text{ then } \begin{cases} \lambda_i^* > 0 \Rightarrow g_i(x^*) = 0 \\ g_i(x^*) < 0 \Rightarrow \lambda_i^* = 0 \end{cases}$$

↳ active constraint at the optimum

Two key corollaries

Corollary 1 (sufficient optimality conditions for the primal)

If x^*, λ^*, μ^* simultaneously verify

(a) $x^* = \arg\min_x L(x, \lambda^*, \mu^*)$

(b) $\lambda_i^* g_i(x^*) = 0 \quad i = 1, \dots, m$

(c) $g(x^*) \leq 0$

(d) $h(x^*) = 0$

$\Big\}$ primal feasibility

(e) $\lambda^* \geq 0$ $\Big\}$ dual feasibility

then, x^* is an optimizer of (P_{min})

Rmk. • (λ^*, μ^*) are a certificate of optimality for x^*
• Alternative way for solving the primal

Proof. $$f(x^*) = f(x^*) + \underbrace{\lambda^{*T} g(x^*)}_{=0 \text{ from point (b)}} + \underbrace{\mu^{*T} h(x^*)}_{=0} \leq$$

$$\leq f(x) + \lambda^{*T} g(x) + \mu^{*T} h(x) \leq f(x)$$

↳ holds because of (a) ↳ holds $\forall x \in X$

We have shown $f(x^*) \leq f(x)$, $\forall x \in X$ ∎

Rmk. · (a) - (e) are called optimality conditions

· The corollary does not assume convexity of (P_{min})

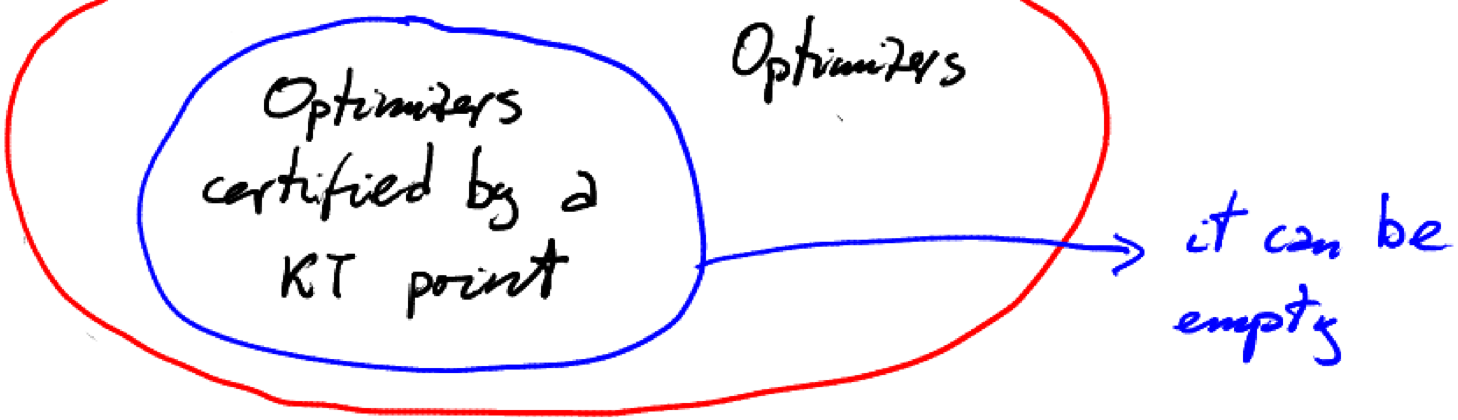
· Strong duality is not required but it is implied

Corollary 2 (necessary optimality conditions for convex programming)

If (P_{min}) is convex and constraints are qualified, then x^* is a primal optimizer only if $\exists\, (\lambda^*, \mu^*)$ verifying the optimality conditions.

Proof. Constraints qualified $\Rightarrow$ there is a KT point. From the Theorem, we have (a) and (b) verified. From primal and dual feasibility also (c), (d) and (e) are verified ∎

Rmk. For general optimization problems one has



For convex programming, the two sets coincide if constraints are qualified

↳ Solving the primal is as difficult as solving the optimality conditions

Karush-Kuhn-Tucker (KKT) conditions

Focus on the optimality condition (a): $x^* = \operatorname{argmin} L(x, \lambda^*, \mu^*)$

Assumption 1. $f, g, h \in C^1$

Then $L \in C^1$ and if x^* is a minimizer, then $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$, i.e.

$$\nabla_x f(x)\Big|_{x=x^*} + {\lambda^*}^T \nabla_x g(x)\Big|_{x=x^*} + {\mu^*}^T \nabla_x h(x)\Big|_{x=x^*} = 0 \qquad (a')$$

Def. The KKT conditions are the optimality conditions with (a) replaced by (a')

Rmk. One can show that if the primal is convex, then

$$(a') \text{ holds} \iff x^* \text{ is a minimizer of } L \in C^1$$

Corollary 1' (sufficiency of KKT)

If (P_{min}) is convex, Assumption 1 holds and x^*, λ^*, μ^* verify the KKT conditions, then x^* is a primal optimizer and (λ^*, μ^*) is a KT point

Corollary 2' (necessity of KKT)

If (P_{min}) is convex, Assumption 1 holds and constraints are qualified, then x^* is an optimizer of (P_{min}) only if $\exists\, (\lambda^*, \mu^*)$ verifying the KKT conditions

Rmk.

- For convex problems with qualified constraints, KKT are necessary and sufficient
- Generalizations of the corollaries to nonconvex problems exist
- Ad hoc algorithms for solving KKT have been developed