Strong Duality via Convex Conjugacy

Aaron Mishkin
amishkin@cs.stanford.edu

Abstract

We develop Lagrangian duality using only convex conjugacy and the convex/concave closure of functions. We relate the Lagrange dual problem to the convex closure of the Lagrangian by introducing a primal perturbation. Then, we show how the duality gap can be expressed as the difference of the convex closure of the Lagrangian and the concave closure of the dual-optimal value function. Conditions for these concave/convex closures to recover the original functions immediately give sufficient conditions for the duality gap to be zero. A symmetric argument using a dual perturbation shows the duality gap can also be characterized by the concave closure of the Lagrangian and convex closure of the primal-optimal value function. This second characterization leads to a simple proof that Slater's condition is sufficient for strong duality attain.

1 Introduction

The topic of our analysis is the inequality constrained optimization problem,

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0. \tag{1}$$

We assume that $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are proper, but not necessarily convex functions. Let $\lambda \in \mathbb{R}^m_+$ be the dual parameters associated with this problem. The Lagrangian is

$$\mathcal{L}(x, \lambda) := f(x) + \langle \lambda, g(x) \rangle . \tag{2}$$

The essential objective incorporates the constraint into the domain of f and can be obtained by maximizing the Lagrangian over λ as follows:

$$f_{\text{ess}}(x) := \sup_{\lambda \geq 0} \mathcal{L}(x, \lambda) = \begin{cases} f(x) & \text{if } g(x) \leq 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{3}$$

Minimizing the essential objective is equivalent to solving Problem 1. We denote the optimal value of the primal problem as $p^\star = \min_x f_{\text{ess}}(x)$.

We will find it useful to define a Lagrangian which incorporates non-negativity of λ into the domain. Let the essential Lagrangian be

$$\bar{\mathcal{L}}(x, \lambda) = \mathcal{L}(x, \lambda) + \mathbb{1}_{\geq 0}(\lambda),$$

which is ∞ when $\lambda \not\geq 0$ and $\mathcal{L}$ otherwise. The Lagrange dual function $d(\lambda)$ is given by minimizing the Lagrangian over x,

$$d(\lambda) := \inf_x \mathcal{L}(x, \lambda). \tag{4}$$

The dual-optimal value is then obtained maximizing the dual function over $\lambda \geq 0$,

$$d^\star = \sup_{\lambda \geq 0} d(\lambda). \tag{5}$$

Weak duality states that $p^\star \geq d^\star$; weak duality always holds by the saddle-point property:

$$p^\star = \inf_x \sup_{\lambda \geq 0} \mathcal{L}(x, \lambda) \geq \sup_{\lambda \geq 0} \inf_x \mathcal{L}(x, \lambda) = d^\star. \tag{6}$$

We say that the *zero duality gap* property holds when $p^\star = d^\star$. *Strong duality* attains when the zero duality gap property holds and the dual problem admits at least one finite solution.

1.1 Technical Preliminaries

Now we introduce background which will be necessary for our analysis. Throughout this section, let $h : \mathbb{R}^d \to \mathbb{R}$ be a proper function, meaning $h(x) > \infty$ for every x and there exists x such that $h(x) < \infty$. The domain of h is denoted $\text{Dom}(h) = \{x : h(x) < \infty\}$. Note that h is proper means $\text{Dom}(h)$ is non-empty.

We say that h is lower semi-continuous at x if for every sequence $\{x_k\}$, $x_k \to x$, it holds that

$$h(x) \leq \liminf_k h(x_k). \tag{7}$$

Clearly h is lower semi-continuous at x if it is continuous at x. We say that h is lower semi-continuous if it is lower semi-continuous at every $x \in \text{Dom}(h)$. Finally, h is called closed if h is lower semi-continuous and $\text{Dom}(h)$ is closed.

Recall that the epigraph of h is the set

$$\text{epi}(h) := \{(x, \alpha) : h(x) \leq \alpha\}.$$

It is not hard to show that $\text{epi}(h)$ is convex if and only if h is convex. The convex closure of a set $\mathcal{C}$ is the closure of the convex hull of $\mathcal{C}$, denoted by $\text{Cl}(\text{Conv}(\mathcal{C}))$. The function

$$\text{Conv}(h)(x) := \inf\{\alpha : (x, \alpha) \in \text{Cl}(\text{Conv}(\mathcal{C}))\},$$

is called the convex closure of h. It is the largest convex function which is majorized by h, meaning $\text{Conv}(h)(x) \leq h(x)$ for all $x \in \text{dom}(h)$. The concave closure of a function is obtain by taking the convex closure of its negative: $\text{Concave}(h) = -\text{Conv}(-h)$. Our analysis will use the convex/concave closure of functions to characterize the duality gap.

One way to compute the closure of a function is through conjugacy. The convex conjugate of h is the function

$$h^*(y) = \sup_x \{\langle y, x \rangle - h(x)\}. \tag{8}$$

The conjugate h^* is the supremum of a collection of affine functions and so is convex in y regardless of convexity of h. Moreover, if $\bar{x}$ achieves the supremum in eq. (8) and h is sub-differentiable at $\bar{x}$, then h^* acts as the inverse subgradient mapping:

$$\bar{x} \in \arg\min_x \{h(x) - \langle y, x \rangle\} \implies y \in \partial h(\bar{x}).$$

If h^* is proper, then the famous Fenchel–Moreau theorem (Bertsekas, 2009a, Proposition 1.6.1) states that the biconjugate

$$h^{**}(y^*) = \sup_y \{\langle y^*, y \rangle - h^*(y)\},$$

is exactly the convex closure operation: $h^{**} = \text{Conv}(h)$. If h is lower semi-continuous and convex, then $\text{Conv}(h) = h = h^{**}$. However, in general h may not be lower semi-continuous at every point in its domain; therefore, we provide a more fine-grained characterization in the following lemma.

Lemma 1. *Suppose h is convex. Then $\text{Conv}\,(h)\,(x) = h(x)$ if and only if h is lower semi-continuous at x.*

Proof. Suppose h is lower semi-continuous at x and let $\{x_k, \alpha_k\} \subset \text{epi}h$ such that $x_k \to x$ and $\alpha_k \to \alpha$. We have

$$f(x) \leq \liminf_k f(x_k) \leq \alpha,$$

which implies $(x, \alpha) \in \text{epi}(f)$. Finally, $f(x) \leq \text{Conv}\,(h)\,(x)$ holds as claimed.

Now, suppose that $\text{Conv}\,(h)\,(x) = h(x)$ but h is not lower semi-continuous at x. Then there exists a sequence $x_k \to x$ such that $\liminf_k f(x_k) < f(x)$. Dropping to a subsequence if necessary, we find that $(x, \liminf_k f(x_k)) \in \text{Cl(epi f)}$ so that

$$\text{Conv}\,(h)\,(x) \leq \liminf_k f(x_k) < h(x),$$

which is a contradiction. □

We conclude this section with a technical result relating sub-differentiability of h^* to minimizers of h.

Lemma 2. *Let h be convex and lower semi-continuous. The solution set to the minimization problem*

$$\inf_x h(x),$$

is exactly $\partial h^*(0)$.

Proof. We start with the observation

$$h^*(0) = \sup_x -h(x) = -\inf_x h(x).$$

Now, suppose that $\bar{x} \in \arg\min_x h(x)$. Then, $h^*(0) = -h(\bar{x})$ and

$$h^*(y) = \sup_x \langle x, y \rangle - h(x) \geq \langle \bar{x}, y \rangle - h(\bar{x}) = h^*(0) + \langle \bar{x}, y - 0 \rangle\,,$$

which shows that $\bar{x} \in \partial h^*(0)$.

For the reverse inclusion, suppose $x \in \partial h^*(0)$. Then,

$$\inf_x h(x) = -h^*(0) \geq \langle \bar{x}, y \rangle - h^*(y),$$

for all $y \in \mathbb{R}^d$. Taking the supremum on the right-hand side implies $\inf_x h(x) \geq h^{**}(\bar{x}) = h(\bar{x})$, where equality holds by Fenchel-Moreau theorem. This completes the proof. □

2 Primal Perturbations and Convex Conjugacy

We start by forming a family of perturbed Lagrangian functions,

$$P(x, \lambda, s) = \mathcal{L}(x, \lambda) - \langle x, s \rangle\,. \tag{9}$$

Adding this perturbation in the primal parameter gives rise to a parameterized family of dual functions,

$$d(\lambda, s) = \inf_x P(x, \lambda, s) \tag{10}$$

as well as a parameterized dual-optimal value function,

$$d^\star(s) = \sup_{\lambda \geq 0} d(\lambda, s). \tag{11}$$

The function $s \mapsto d^\star(s)$ measures sensitivity of the dual-optimal value to linear tilts of the primal optimization problem. When there is no perturbation, $d^\star(0) = d^\star$. The parameterized dual function and dual-optimal value function behave much like the standard versions.

Lemma 3. *The parameterized Lagrange dual function $(s, \lambda) \mapsto d(s, \lambda)$ is concave and the dual-optimal value function $d^\star$ is also concave.*

Proof. The perturbed Lagrangian P is affine in both s and λ, so that $d(\lambda, s)$ is obtained by minimizing a collection of affine functions. As a result, $d(\lambda, s)$ is jointly concave in s and λ. Partial maximization of a concave function preserves concavity, from which we deduce $d^\star$ is a concave function. □

Now we view the parameterized dual problem through the lens of conjugacy. The perturbed Lagrangian $d(\lambda, s)$ has another interpretation as the conjugate of the Lagrangian function,

$$d(\lambda, s) = \inf_x \mathcal{L}(x, \lambda) - \langle x, s \rangle = -\mathcal{L}^*(s, \lambda).$$

This reveals a connection between $\mathcal{L}$ and the conjugate of $d(\lambda, s)$: the conjugate of $-d$ is the convex closure of $\mathcal{L}$,

$$(-d(\lambda, \cdot))^*(y) = \sup_s \langle y, s \rangle - \mathcal{L}^*(s, \lambda) = \mathcal{L}^{**}(y, \lambda) = \text{Conv}(\mathcal{L})(y, \lambda).$$

As a result, the bi-conjugate $\mathcal{L}^{**}$ is both convex and closed and $\mathcal{L}^{**} \leq \mathcal{L}$. Thus, we can lower-bound the duality gap by controlling the duality gap for the closure $\mathcal{L}^{**}$. We start by maximizing $\mathcal{L}^{**}$ to obtain a lower bound on the essential objective.

Lemma 4. *The essential objective satisfies*

$$f_{ess}(x, \lambda) \geq \sup_{\lambda \geq 0} \mathcal{L}^{**}(x, \lambda) = (-d^\star)^*(y), \tag{12}$$

where $d^\star$ is the dual-optimal value function.

Proof. Maximizing $\mathcal{L}^{**}$ immediately yields the first lower bound on the essential objective:

$$\sup_{\lambda \geq 0} \mathcal{L}^{**}(x, \lambda) \leq \sup_{\lambda \geq 0} \mathcal{L}(x, \lambda) = f_{\text{ess}}(x, \lambda).$$

Our goal is write the left-hand side in terms of the dual function, since this latter object is the "natural" lower-bound on the essential objective stemming from weak duality. We do this as follows:

$$\begin{aligned}
\sup_{\lambda \geq 0} \mathcal{L}^{**}(y, \lambda) &= \sup_{\lambda \geq 0} \sup_s \left\{ \langle y, s \rangle - \mathcal{L}^*(s, \lambda) \right\} \\
&= \sup_{\lambda \geq 0} \sup_s \left\{ \langle y, s \rangle + d(\lambda, s) \right\} \\
&= \sup_s \langle y, s \rangle + \sup_{\lambda \geq 0} \left\{ d(\lambda, s) \right\} \\
&= (-\sup_{\lambda \geq 0} d(\lambda, \cdot))^*(y) \\
&= (-d^\star)^*(y),
\end{aligned}$$

where $d^\star$ is the dual-optimal value function. □

We have uncovered another connection with conjugacy. We can also see that minimizing the left-hand side of eq. (12) immediately gives a saddle-point equation much like that required for the zero duality gap property. Since the right-hand side is a convex conjugate, it is natural to write minimization as another conjugacy operation:

$$\begin{aligned}\inf_y \sup_{\lambda \geq 0} \mathcal{L}^{**}(y, \lambda) &= -(\sup_{\lambda \geq 0} \mathcal{L}^{**}(\cdot, \lambda))^*(0) \\ &= -(-d^\star)^{**}(0).\end{aligned}$$

Recalling $d^\star(0) = d^\star$ shows we have obtained a zero duality gap result for the inf-sup value of $\text{Conv}(\mathcal{L})$ and the concave closure of the dual-optimal value function.

Proposition 1. *The duality gap is exactly characterized by closure operations on the Lagrangian and the dual-optimal value function. In particular,*

$$\delta_{gap} = [p^\star - \inf_y \sup_{\lambda \geq 0} Conv(\mathcal{L})(y, \lambda)] + [Concave(d^\star)(0) - d^\star]. \tag{13}$$

The proof follows immediately from our discussion above. As result, there can be zero duality gap when the convex/concave closures match the original functions.

Proposition 2. *The zero duality gap property holds if f is convex and closed and if $s \mapsto d^\star(s)$ is lower semi-continuous at* 0*.*

Proof. If f is convex and closed, then $x \mapsto \mathcal{L}(x, \lambda)$ is convex and closed for each λ and $\mathcal{L}^{**}(\cdot, \lambda) = \mathcal{L}(\cdot, \lambda)$. We obtain

$$\inf_x \sup_{\lambda \geq 0} \mathcal{L}^{**}(y, \lambda) = -(-d^\star)^{**}(0), \tag{14}$$

so that the duality gap can come only comes from the concave closure of $d^\star$. Since $d^\star$ is a concave function, we know from lemma 1 that

$$-(-d^\star)^{**}(0) = d^\star(0)$$

if and only if $d^\star$ is lower semi-continuous at 0. This completes the proof. □

Since $s \mapsto d^\star(s)$ is concave, it is continuous on the interior of its domain. Thus, $d^\star(s) > -\infty$ on a neighbourhood $\mathcal{N}$ containing 0 is a sufficient condition for zero duality gap.

3 Dual Perturbations and Convex Conjugacy

The previous section showed that we can characterize the duality gap in terms of the convex closure of $\mathcal{L}$ and the concave closure of the dual-optimal value function. This arose through the introduction of a primal perturbation parameter, s. However, it is difficult to connect $d^\star$ to constraint qualifications, such as Slater's condition, because the interaction between s and the constraint set is subtle.

In this section, we avoid such difficulties by consider perturbations in the dual space. In particular, we consider perturbations of the constraint,

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq -r, \tag{15}$$

where r is a dual perturbation vector. We call this a dual perturbation because it leads to a linear tilt of the Lagrangian in the dual parameter,

$$D(x, \lambda, r) = \mathcal{L}(x, \lambda) + \langle \lambda, r \rangle \tag{16}$$

The analysis proceeds now similarly to the primal perturbation; maximizing $D(x, \lambda, r)$ in λ to obtain a perturbed essential objective gives

$$\begin{aligned} f_{\text{ess}}(x, r) &= \sup_{\lambda \geq 0} D(x, \lambda, r) \\ &= \sup_{\lambda} \mathcal{L}(x, \lambda) + \langle \lambda, r \rangle + \mathbb{1}_{\geq 0}(\lambda) \\ &= \sup_{\lambda} \bar{\mathcal{L}}(x, \lambda) + \langle \lambda, r \rangle \\ &= (-\bar{\mathcal{L}}(x, \cdot))^*(r), \end{aligned}$$

where we recall $\text{Dom}(\bar{\mathcal{L}}) = \text{Dom}(f) \times \mathbb{R}^m_+$. The conjugate of $f_{\text{ess}}(x, r)$ is thus related to the concave closure of $\mathcal{L}$ as follows:

$$-f^*_{\text{ess}}(x, r^*) = -(-\bar{\mathcal{L}}(x, \cdot))^{**}(r^*).$$

Minimizing both sides with respect to x recovers the conjugate of the primal-optimal value function,

$$\begin{aligned} \inf_x -f^*_{\text{ess}}(x, r^*) &= -\sup_x \sup_r \{\langle r, r^* \rangle - f_{\text{ess}}(x, r)\} \\ &= -\sup_r \left\{\langle r, r^* \rangle - \inf_x f_{\text{ess}}(x, r)\right\} \\ &= -\sup_r \{\langle r, r^* \rangle - p^\star(r)\} \\ &= -(p^\star)^*(r^*). \end{aligned}$$

That is,

$$(p^\star)^*(r^*) = \sup_x (-\bar{\mathcal{L}}(x, \cdot))^{**}(r^*).$$

Taking conjugates once more and evaluating at $r = 0$, we obtain

$$(p^\star)^{**}(0) = \sup_{r^*} \inf_x -(-\bar{\mathcal{L}}(x, \cdot))^{**}(r^*).$$

Proposition 3. *The duality gap admits a second characterization in terms of of convex/concave closure operations. Specifically,*

$$\delta_{gap} = [p^\star - Conv(p^\star)(0)] + [\sup_r \inf_x Concave(\mathcal{L})^{**}(x, r^*) - d^\star]. \tag{17}$$

As a result, the zero duality gap property holds when f and g are convex and closed and the primal-optimal value function $p^\star$ is lower semi-continuous at 0.

Proof. The first part of the claim follows from the calculations above. For the second part, we first consider the concave closure of $\bar{\mathcal{L}}$ in λ. By definition,

$$\mathcal{L}(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle + \mathbb{1}_{\geq 0}(\lambda).$$

This is an affine function in λ with closed, convex domain $\mathbb{R}^m_+$, so that $\lambda \mapsto -\mathcal{L}(x, \lambda)$ is closed and convex for each x. Thus, $-(-\bar{\mathcal{L}})^{**}(x, \dot{)} = \mathcal{L}(x, \cdot)$.

Now we turn to the primal-optimal value function. Observe that

$$f_{\text{ess}}(x, r) = \sup_{\lambda \geq 0} D(x, \lambda, r) = f(x) + \mathbb{1}_{\leq r}(g(x)),$$

and $p^\star(r) = \inf_x f_{\text{ess}}(x, r)$ is obtained by partial minimization of a convex function and thus is convex. Partial minimization does not necessary preserve closedness, meaning $p^\star(r)$ may not be closed everywhere. However, by assumption $p^\star$ is lower semi-continuous at 0 and thus $(p^\star)^{**}(0) = p^\star(0)$. This completes the proof. □

4 Existence of Solutions and Strong Duality

Now we address sufficient conditions for strong duality to hold. Recall that strong duality requires both the zero duality gap property and existence at least one dual solution. Let $\mathcal{P}(s)$ and $\mathcal{D}(r)$ denote the primal and dual solution sets as functions of the primal/dual perturbations, respectively. We start by showing the existence of primal solutions via sub-differentiability of the dual-optimal value function.

Lemma 5. *Suppose f and g are convex and closed. If Concave($d^\star$) is sub-differentiable at 0, then*

$$\mathcal{P}(0) = -\partial \textit{Concave}(d^\star)(0).$$

Similarly, if Conv($p^\star$) is sub-differentiable at 0, then

$$\mathcal{D}(0) = -\partial \textit{Conv}(p^\star)(0).$$

Proof. Recall that $(-d^\star)^* = \sup_{\lambda \geq 0} \mathcal{L}^{**}(x, \lambda)$. Since f, g are convex and closed, $\mathcal{L}^{**} = \mathcal{L}$ and $(-d^\star)^* = f_{\text{ess}}$. Partial maximization preserves closedness (Bertsekas, 2009a, Proposition 1.1.6), from which we deduce that f_{ess} is closed (alternatively, this follows since $g(x) \leq 0$ defines a closed set). Now it remains only to invoke lemma 2 to see that

$$\partial(-d^\star)^{**}(0) = \arg\min_x f_{\text{ess}}(x),$$

completing this part of the proof.

The second part of the proof follows by a symmetric argument. Recall that

$$p^\star)^*(r^*) = \sup_x (-\bar{\mathcal{L}}(x, \cdot))^{**}(r^*) = \sup_x -\bar{\mathcal{L}}(x, r^*),$$

where the right-hand is closed and convex as the supremum of affine functions. Invoking lemma 2 gives

$$\partial(p^\star)^{**}(0) = \arg\min_\lambda \sup_x -\bar{\mathcal{L}}(x, \lambda) = -\arg\max_{\lambda \geq 0} \inf_x \mathcal{L}(x, \lambda),$$

which completes the proof. □

Taken together, proposition 3 and lemma 5 imply that strong duality holds whenever $p^\star$ is lower semi-continuous and sub-differentiable at 0. We will show that the following condition, called Slater's constraint qualification, full-fills both of these requirements.

Assumption 1 (Slater's Constraint Qualification)**.** *There exists a strictly feasible point for Problem 1. That is, there exists $\bar{x} \in$ Dom(f) such that $g(\bar{x}) < 0$.*

Critically, Slater's condition ensures that the primal-optimal value is well-behaved with respect to dual perturbations on a neighbourhood around 0.

Lemma 6. *Assume $p^\star < \infty$. If Slater's constraint qualification holds, then the primal-optimal value function is continuous on a neighbourhood containing 0.*

Proof. Since $r \mapsto p^\star(r)$ is convex and all convex functions are continuous on the interior of their domains, it is sufficient to show that 0 is in the interior of $\text{Dom}(p)$. Thus, we must show $p(r) < \infty$ for all r in some neighbourhood of 0.

Start by observing that $r \leq 0$ relaxes the constraints, implying

$$p^\star(r) = \min_x \{f(x) : g(x) \leq r\} \leq p^\star(0) < \infty,$$

where the last inequality holds by assumption. On the other hand, if $0 < r \leq -g(\bar{x})$, then

$$\begin{aligned} p^\star(r) &\leq f_{\text{ess}}(\bar{x}, r) \\ &= \sup_{\lambda \geq 0} f(\bar{x}) + \langle \lambda, g(x) + r \rangle \\ &= f(\bar{x}) < \infty. \end{aligned}$$

Since $g(\bar{x}) < 0$, we have constructed a neighbourhood on which $p^\star(r)$ is finite. This completes the proof. □

We can now prove that Slater's constraint qualification is sufficient for strong duality.

Proposition 4. *Assume $p^\star < \infty$ and that Slater's constraint qualification holds. Then strong duality attains and the set of dual solutions is compact.*

Proof. Since $p^\star$ is continuous at 0 (lemma 6), it follows that $p^\star$ is lower semi-continuous at 0 and lemma 1 suffices to show $(p^\star)^{**}(0) = p^*$. The zero duality gap property now follows immediately from the dual-perturbation characterization of the duality gap in proposition 3.

Continuity of $p^\star$ at 0 and convexity of $(p^\star)^{**}$ guarantee that $(p^\star)^{**}$ is sub-differentiable at 0. lemma 5 now establishes the set of dual solutions to be $-\partial(p^\star)^{**}(0)$; In particular, $\partial(p^\star)^{**}(0)$ is non-empty and compact since 0 is in the interior of $\text{Dom}((p^\star)^{**})$ (Bertsekas, 2009a, Proposition 5.4.1). □

5 Further Reading

This note was inspired by an analysis of semi-infinite optimization problems in the review by Shapiro (2009). The approach to deriving sufficient conditions for zero duality gap using primal perturbations in section 2 is adapted from this work, although we develop the argument in more detail and with an emphasis on convex/concave closures. Another difference is that we deal only with a finite number of constraints, while Shapiro considers $g(x, \omega) \leq 0$ for all $\omega \in \Omega$, where Ω is a potentially infinite index set.

Rockafellar (1974) covers the connection between conjugacy and duality theory in both finite and infinite spaces. Their work also focuses on the importance of convex/concave closures in describing, although this is part of a much wider development, rather than specifically in the context of the duality gap.

Finally, Bertsekas (2009b) provides an alternative and complementary approach to deriving strong duality. They use dual perturbations to establish a geometric duality framework, where the duality gap is described by the gap between supporting hyperplanes to $\text{epi}(p^\star)$ and $p^\star(0)$. Our reliance on the conjugate operation implicitly captures this geometry, with the same gap arising in our work as the difference between $p^\star$ and $(p^\star)^{**}$. Furthermore, while Bertsekas (2009b) guarantee strong duality directly through existence of a non-vertical separating hyperplane, we avoid geometry arguments through the use of Fenchel-Moreau theorem.

References

[Ber09a] Dimitri Bertsekas. *Convex optimization theory*. Vol. 1. Athena Scientific, 2009.

[Ber09b] Dimitri Bertsekas. "Min common/max crossing duality: A geometric view of conjugacy in convex optimization". In: *Lab. for Information and Decision Systems, MIT, Tech. Rep. Report LIDS-P-2796* (2009).

[Roc74] R Tyrrell Rockafellar. *Conjugate duality and optimization*. SIAM, 1974.

[Sha09] Alexander Shapiro. "Semi-infinite programming, duality, discretization and optimality conditions". In: *Optimization* 58.2 (2009), pp. 133–161.