# LECTURE 5

## DUALITY

### LAGRANGIAN DUALITY

- optimization problems in standard form
- Lagrangian $\mathcal{L}$, dual function $d$, dual problem
- Constraint qualification (e.g. Slater) involves dom and explicit constraints
- KKT optimality conditions

### FENCHEL DUALITY

- composite optimization problems
- function $\mathcal{L}$, min-max inequality, (Fenchel) dual problem
- Constraint qualification only involves dom (only implicit constraints appear in Fenchel duality)
- Optimality conditions expressed in terms of subdifferentials (Fermat's rule, Fenchel-Young's equality)

# LAGRANGIAN DUALITY – standard problem

PRIMAL

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \le 0 \quad i = 1, \dots, m \\ & h_j(x) = 0 \quad j = 1, \dots, p \end{aligned}$$

$x \in \mathbb{R}^n$

$D \subseteq \mathbb{R}^n$ is the domain: set of $x$ for which the objective and all constraint functions are defined

$p^*$ optimal value

LAGRANGIAN $\mathcal{L} : D \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$\mathcal{L}(x, \lambda, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \nu_j h_j(x)$$

LAGRANGE DUAL FUNCTION $d : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$d(\lambda, \nu) = \inf_{x \in D} \mathcal{L}(x, \lambda, \nu)$$

$d$ is pointwise infimum of a family of affine functions $\to$ concave

LOWER BOUND PROPERTY

If $\lambda \geq 0$ (entrywise)

$$d(\lambda, \nu) \le p^*$$

LAGRANGE DUAL PROBLEM

DUAL

$$\begin{aligned} d^* = \text{maximize} \quad & d(\lambda, \nu) \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned}$$

## WEAK DUALITY (always holds)

$$d^\star \leq p^\star$$

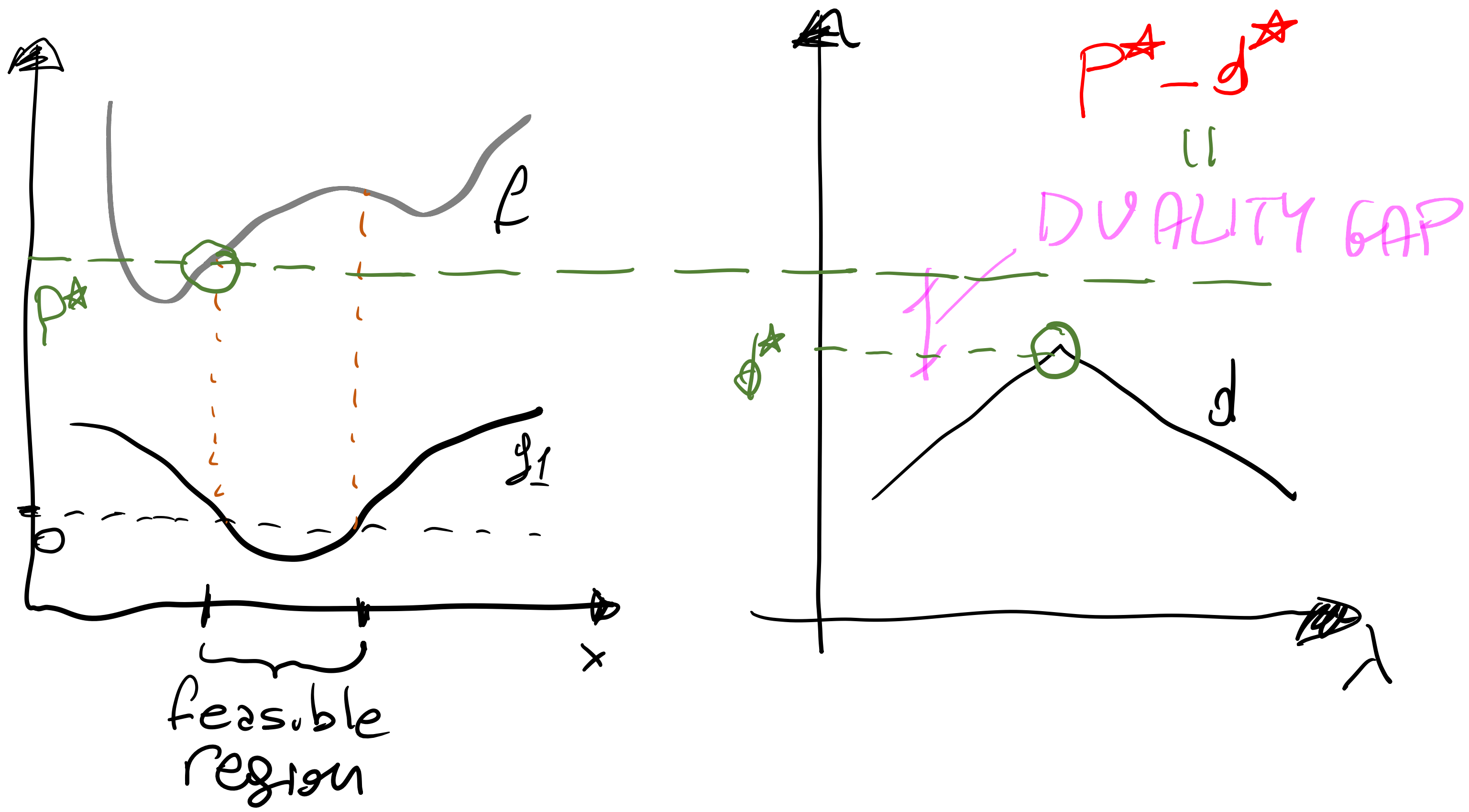
## STRONG DUALITY (certified by constraint qualification)

$$d^\star = p^\star$$

### Example 1

$m = 1,\ p = 0$

$f, g_1$ non convex



## Example 2 — Orthogonal projection onto the Unit-Simplex

$$\begin{aligned} \min \ & \|x-y\|_2^2 \\ \text{s.t.} \ & \mathbf{1}^T x = 1 \\ & x \geq 0 \end{aligned}$$

$y \in \mathbb{R}^n$ is the given vector we want to project

$$\mathcal{L}(x,\lambda) = \|x-y\|_2^2 + 2\lambda(\mathbf{1}^T x - 1) = \|x\|_2^2 - 2(y-\lambda\mathbf{1})^T x +$$

$$+\|y\|_2^2 - 2\lambda = \underbrace{\sum_{j=1}^{n}\left(x_j^2 - 2(y_j-\lambda)x_j\right)}_{\text{only this part depends on } x} + \|y\|_2^2 - 2\lambda$$

Note: we don't dualize the inequality constraint and we consider it as implicit constraint in the dual problem

Only this part depends on $x$. Moreover, it depends separately on $x_j$

n separate problems:

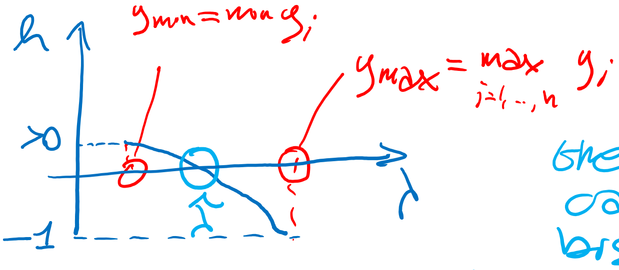
$$\min_{x_j \geq 0} \ x_j^2 - 2(y_j-\lambda)x_j \quad \begin{cases} \text{minimizer } x^* = \begin{cases} y_j - \lambda, & y_j \geq \lambda \\ 0, & \text{else} \end{cases} \triangleq [y_j-\lambda]_+ \\ \text{optimal value: } -[y_j-\lambda]_+^2 \end{cases}$$

The dual problem is:

$$\max_{\lambda} \ \underbrace{-\sum_{j=1}^{n}[y_j-\lambda]_+^2 - 2\lambda + \|y\|_2^2}_{d(\lambda)}$$

This problem can be solved by bisection and. Note:

- $d$ by definition is concave and $\lim_{\lambda\to-\infty} d = \lim_{\lambda\to\infty} d = -\infty$
- the optimal $\lambda$ is such that $d'(\lambda)=0 \to \underbrace{\sum_{j=1}^{n}[y_j-\lambda]_+ - 1}_{h(\lambda)} = 0$
- $h$ has this trend




the crossing point $\tilde{\lambda}$ can be found by bisection

$$d^* = d(\tilde{\lambda})$$

# CONSTRAINT QUALIFICATION (CQ)

very useful to infer if STRONG DUALITY holds.

One example is SLATER'S CONDITION:

- For a convex problem, strong duality holds if

$$\exists x \in \text{int } D \text{ s.t. } \left.\begin{array}{l} g_i(x) < 0 \ \forall i \\ Ax = b \end{array}\right\}$$ i.e. if the problem is strictly feasible

- Sharpened Slater

$$\exists x \in \text{relint } D \text{ s.t. } \begin{array}{ll} g_i(x) \le 0 & i = 1, \dots, k \\ g_i(x) < 0 & i = k+1, \dots m \end{array}$$ affine inequalities

That is, if $x \in$ relint $D$, then affine inequalities can also be satisfied non-strictly.

these are the only ones in LP and QP

Slater condition reduces to feasibility

**Example 3**

$$D = \{(x,y) \mid y > 0\} \qquad \min e^{-x} \quad \text{s.t. } \frac{x^2}{y} \le 0$$

$f(x,y)$

$$\nabla^2 f(x,y) = \frac{2}{y^3}\begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} = \frac{2}{y^3}\begin{bmatrix} y \\ -x \end{bmatrix}\begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$$

The problem is convex.

Does Slater condition hold?

The problem is not strictly feasible since the only feasible $x$ is $x = 0$ (with any $y$)

Therefore, Slater condition is not satisfied and nothing can be said on strong duality.

DUAL

$$d(\lambda) = \inf\left(e^{-x} + \lambda \frac{x^2}{y}\right) = \begin{cases} 0 & \lambda \ge 0 \ * \\ -\infty & \lambda < 0 \end{cases}$$

$d^* = 0$

$p^* = 1 \rightarrow \boxed{p^* - d^* = 1}$

* Set $x = t$, $y = t^3$

for $t \to \infty$, $\lambda \ge 0$ the objective function $\to 0$

# KARUSH-KUHN-TUCKER (KKT) CONDITIONS

They are conditions on the primal ($x$) and dual ($\lambda, \nu$) variables.

We make the assumption that $f, g_i, h_i$ are differentiable (not necessarily convex).

Assume $p^* = d^*$ and let $x^*, \lambda^*, \nu^*$ the associated primal and dual optimal solutions.

Then, they satisfy:

KKT
1. Primal feasibility
2. Dual feasibility
3. Complementary slackness
4. stationarity ($\nabla_x \mathcal{L} = 0$)

- If the primal is convex, KKT are also sufficient
- If the primal is convex and satisfies CQ, then the duality gap is zero, so $x$ is optimal iff there are $\lambda, \nu$ satisfying KKT $\Rightarrow$ they provide necessary and sufficient conditions for optimality.

$\Downarrow$

Some optimization algorithms are essentially efficient methods for solving the KKT conditions (but for non-convex problems these are only necessary)

# FENCHEL DUALITY – composite problem

$$\min f(x) + g(Lx)$$

$f, g$ convex, $L$ linear operator (in these notes, a matrix)

- no explicit constraints (indicator function used as proxy)
- differentiability not required

PRIMAL → an inf sup formulation $\mathcal{L}(x,y,z)$

$$\underbrace{\inf_{x,y}}_{\text{This is the standard primal minimization}} \; \underbrace{\sup_{\mu}}_{\text{This enforces the constraint } Lx=y} \left\{ f(x) + g(y) + \mu^T(Lx - y) \right\}$$

This is the standard primal minimization

This enforces the constraint $Lx = y$

## Minimax inequality

obtained swapping sup/inf. It always holds.

$$p^* = \underbrace{\inf_{x,y} \sup_{\mu} \mathcal{L}(x,y,\mu)}_{} \geq \underbrace{\sup_{\mu} \inf_{x,y} \mathcal{L}(x,y,z)}_{} = d^*$$

Fenchel primal problem and dual problem

It can be expressed in terms of conjugate functions.

$\boxed{\text{Weak duality}}$ → $p^* \geq d^*$

$$\inf_{x,y} f(x) + g(Lx) \geq \sup_{\mu} -f^*(-L^T\mu) - g^*(\mu)$$

$\boxed{\text{Strong duality}}$ $\quad p^* = d^*$

CQ for Fenchel duality: $f, g$ proper closed convex +

$$\operatorname{relint} \operatorname{dom} f \cap \operatorname{relint} \operatorname{dom}(g \circ L) \neq \emptyset$$

## Example 4

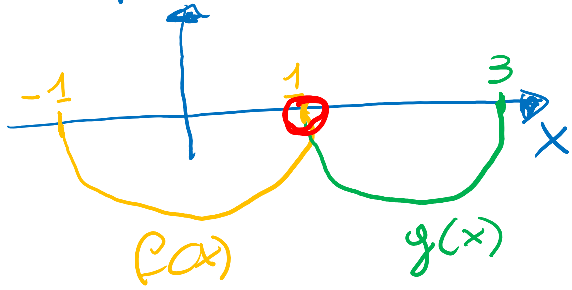
$$f(x) = \begin{cases} -\sqrt{1-x^2}, & -1 \le x \le 1 \\ +\infty, & \text{else} \end{cases}$$

$$L = I, \quad g(x) = \begin{cases} -\sqrt{1-(x-2)^2}, & 1 \le x \le 3 \\ +\infty, & \text{else} \end{cases}$$

$$\min f(x) + g(x)$$

① Let's start by plotting the functions

Both $f$ and $g$ are convex, closed and proper.




$$f(x) + g(x) = \begin{cases} 0, & x = 1 \\ +\infty, & \text{else} \end{cases}$$

$\Rightarrow$ $\begin{cases} p^* = 0 \text{ is the optimal value.} \\ x^* = 1 \text{ is the minimizer for the primal.} \end{cases}$

② The CQ condition is not satisfied.

This comes from the fact that the intersection between the two domains is at a boundary point (indeed $x = 1$).

Let's try to solve the dual

③ Consider $h(x) = -\sqrt{1-(x-a)^2}$

$f(x) = h(x)\big|_{a=0}$

$g(x) = h(x)\big|_{a=2}$

$$h^*(y) = \sup_x \left\{ yx + \sqrt{1-(x-a)^2} \right\}$$

To find a stationary point we set the derivative to 0

$$y + \frac{-2(x-a)}{2\sqrt{1-(x-a)^2}} = 0$$

$$y = \frac{x-a}{\sqrt{1-(x-a)^2}}$$

$$y^2 = \frac{(x-a)^2}{1-(x-a)^2}$$

$$(x-a)^2(1+y^2) = y^2$$

this gives the maximizer

$$(x-a)^2 = \frac{y^2}{1+y^2} \to x = a \pm \frac{y}{\sqrt{1+y^2}}$$

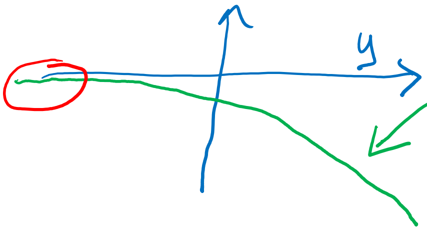
$$h^*(y) = y\left(a + \frac{y}{\sqrt{1+y^2}}\right) + \sqrt{1 - \frac{y^2}{1+y^2}} = ay + \sqrt{1+y^2}$$

$$f^*(y) = \sqrt{1+y^2} \qquad g^*(y) = 2y + \sqrt{1+y^2}$$

Dual problem

$$\max \; -f^*(-\mu) - g^*(\mu)$$

$$\max \; -2\sqrt{1+y^2} - 2y$$




The supremum of the dual function is 0 $\Rightarrow d^* = 0 \; (y \to -\infty)$

This shows that strong duality might hold even if CQ does not hold. Moreover, if CQ is not satisfied, a solution to the dual may not exist.

# Minimax inequality

Given $f(x,y)$, it always holds:

$$\inf_x \sup_y f(x,y) \geq \sup_y \underbrace{\inf_x f(x,y)}_{h(y)}$$

## Proof

$$h(y) \leq f(x,y) \quad (\forall x, y, \text{ by definition})$$

$$\Rightarrow \sup_y h(y) \leq \sup_y f(x,y) \quad (\forall x)$$

$$\Rightarrow \sup_y h(y) \leq \inf_x \sup_y f(x,y)$$

$$\Rightarrow \sup_y \inf_x f(x,y) \leq \inf_x \sup_y f(x,y)$$

$\boxtimes$

Lastly, we need to derive optimality conditions for composite minimization. Unlike KKT, these conditions will need to handle non-differentiable functions.

Useful results:

**FERMAT'S RULE** → generalization of optimality conditions for non-differentiable f

$f: \mathbb{R}^n \to \mathbb{R}$. $x$ minimizes $f$ iff

$$0 \in \underbrace{\partial f(x)}_{\text{subdifferential of } f \text{ at } x}$$

**FENCHEL-YOUNG'S EQUALITY**

$$f^*(y) = y^T x - f(x) \quad \text{iff} \quad y \in \partial f(x)$$

It comes from definition of conjugate function + Fermat

**Examples**

☑ Show that $s \in \partial f(x) \to x \in \partial f^*(s)$
(no assumptions on $f$)

We know that $f(x) \geq f^{**}(x)$

for Fenchel-Young + Fermat

$$0 \overset{!}{=} f^*(s) + f(x) - s^T x \geq f^*(s) + f^{**}(x) - s^T x \geq 0$$

it follows from the definition of conjugate function (also known as Fenchel Young inequality)

Hence, these must all hold with equality.
In particular,

$$0 = f^*(s) + (f^*)^*(x) - s^T x$$

which is equivalent to say $x \in \partial f^*(s)$

☑ Show that $x \in \partial f^*(s) \to s \in \partial f^{**}(x)$

It follows the previous procedure, but now considering $f^*$ instead of $f$.

☑ If $f$ is closed and convex, then
$$f = f^{**}$$
and it is easy to see that iff holds.

This result is known as

## SUBDIFFERENTIAL INVERSE

If $f$ is closed and convex
$$s \in \partial f(x) \iff x \in \partial f^*(s)$$

# OPTIMALITY CONDITIONS

$f, g$ closed convex and CQ holds. Then $x$ solves
$$\min \quad f(x) + g(Lx)$$
iff
$$0 \in \partial f(x) + L^T \partial g(Lx)$$

Obtained from Fermat rule and subdifferential calculus

▪ Primal-dual iff optimality conditions can be obtained using the previously reviewed subdifferential properties

| | |
|---|---|
| $-L^T\mu \in \partial f(x)$ <br> $\mu \in \partial g(Lx)$ | $x \in \partial f^*(-L^T\mu)$ <br> $\mu \in \partial g(Lx)$ |
| $-L^T\mu \in \partial f(x)$ <br> $Lx \in \partial g^*(\mu)$ | $x \in \partial f^*(-L^T\mu)$ <br> $Lx \in \partial g^*(\mu)$ |

subdiff inverse

equivalent to Fermat's rule for the dual problem

#