# Duality, Gradient-free, in Application

Axel Böhm

January 21, 2022

1. Duality

# Duality

Establishes some relation between two classes of objects.

**Definition ("Legendre transform" or "Fenchel conjugate")**

Given a function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we define its **conjugate** $f^* : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ by

$$f^*(y) = \sup_x \{\langle y, x\rangle - f(x)\}$$

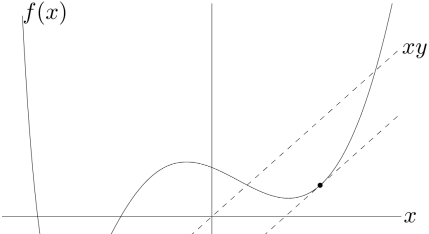


Figure: maximum gap between lin-

# Properties

- $f^*$ is always convex. point-wise max of affine function
- Fenchel's inequality:

$$f(x) + f^*(y) \geq \langle x, y \rangle.$$

- Hence the biconjugate $f^{**} := (f^*)^*$ satisfies $f^{**} \leq f$.
- If f is convex an lsc. then $f^{**} = f$.
- etc.

# Examples

◇ Norm: If $f(x) = \|x\|$, then

$$f^*(y) = \mathbb{1}(\|y\|_* \leq 1),$$

i.e. the indicator of the dual norm ball. Recall the definition of the dual norm:

$$\|y\|_* := \max_{\|x\| \leq 1} \{\langle y, x \rangle\}.$$

In particular: $\|\cdot\|_1 \leftrightarrow \|\cdot\|_\infty$

# More examples

## Generalized linear models

$$\min_{x \in \mathbb{R}^d} f(Ax) + g(x).$$

Two approaches to reformulate:

$$\min_x \max_y \langle y, Ax \rangle - f^*(y) + g(x)$$

## Generalized linear models continued

Or reformulate

$$\min_{x\in\mathbb{R}^d} f(Ax) + g(x)$$

as

$$\min_{x\in\mathbb{R}^d, w\in\mathbb{R}^m} f(w) + g(x) \quad \text{s.t.} w = Ax$$

Use Lagrange function

$$\mathcal{L}(x, w, u) := f(w) + g(x)\langle u, w - Ax\rangle$$

the dual function is given by

$$\varphi(u) = \min_{x\in\mathbb{R}^d, w\in\mathbb{R}^m} \mathcal{L}(x, w, u)$$

Dual problem

$$\max_{u\in\mathbb{R}^m} \varphi(u)$$

## Example: Lasso

$\ell_1$ regularized regression

$$\min_{x \in \mathbb{R}^d} \frac{1}{2}\|Ax - b\|^2 + \lambda \|x\|_1$$

fits this template with

$$f(w) = \frac{1}{2}\|w - b\|^2 \quad \text{and} \quad g(x) = \lambda\|x\|_1$$

Computation gives

$$f^*(u) = \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b - u\|^2 \quad \text{and} \quad g^*(v) = \mathbb{1}(\|v/\lambda\|_\infty \leq 1).$$

# Dual of Lasso

We had

$$f^*(u) = \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b-u\|^2 \quad \text{and} \quad g^*(v) = \mathbb{1}(\|v/\lambda\|_\infty \leq 1).$$

So the dual is

$$\begin{aligned} &\max_{u\in\mathbb{R}^m} -f^*(-u) - g^*(A^T u) \\ \Leftrightarrow &\min_{u\in\mathbb{R}^m} \|b+u\|^2 \quad \text{s.t.} \quad \|A^T u\|_\infty \leq \lambda \end{aligned}$$