Numerical Optimization

CHAPTER 14. DUALITY

TU Dortmund, Dr. Sangkyun Lee

Constrained Optimization

$$\inf_{x \in \mathbb{R}^n} \; f(x)$$

$$\text{s.t.} \quad h_i(x) \geq 0, \; i = 1, \ldots, m$$

$$\ell_j(x) = 0, \; j = 1, \ldots, r$$

➔ Constraint set C

$$\mathcal{D} = \operatorname{dom} f \cap \bigcap_{i=1}^{m} \operatorname{dom} h_i \cap \bigcap_{j=1}^{r} \operatorname{dom} \ell_j \neq \emptyset$$

The functions are not necessarily convex

Constrained Optimization

$$\inf_{x \in \mathbb{R}^n} \; f(x)$$

$$\text{s.t.} \quad h_i(x) \geq 0, \; i = 1, \ldots, m$$
$$\ell_j(x) = 0, \; j = 1, \ldots, r$$

➔ Constraint set C

All functions are smooth, and could be non-convex

The **Lagrangian** function:

$$\mathcal{L}(x; \alpha, \beta) = f(x) - \sum_{i=1}^{m} \alpha_i h_i(x) - \sum_{j=1}^{r} \beta_j \ell_j(x)$$

with Lagrange multipliers $\alpha \in \mathbb{R}^m_+$ and $\beta \in \mathbb{R}^r$.

(implicitly, we define $\mathcal{L}(x; \alpha, \beta) = -\infty$ when $\alpha \not\geq 0$)

Lagrangian forms a lower bound

For any $\alpha \geq 0$ and β (i.e., dual feasible),

$$f(x) \geq \mathcal{L}(x;\alpha,\beta) \quad \text{at each (primal) feasible } x.$$

Obviously, from the definition

$$\mathcal{L}(x;\alpha,\beta) = f(x) - \sum_{i=1}^{m} \underbrace{\alpha_i}_{\geq 0} \underbrace{h_i(x)}_{\geq 0} - \sum_{j=1}^{r} \beta_j \underbrace{\ell_j(x)}_{=0} \leq f(x)$$

Let f^* be the optimal obj. value and C the primal feasible set.

For any $\alpha \geq 0$ and β,

$$f^* \geq \min_{x \in C} \mathcal{L}(x;\alpha,\beta)$$

Dual Objective Function

Let f^* be the optimal obj. value and C the primal feasible set.

For any $\alpha \geq 0$ and β,

$$f^* \geq \inf_{x \in C} \mathcal{L}(x; \alpha, \beta) \geq \inf_{x \in \mathbb{R}^n} \mathcal{L}(x; \alpha, \beta) =: g(\alpha, \beta)$$

$g(\alpha, \beta)$ is the **dual objective function**, which gives a lower bound of f* for any dual feasible u & v.

Dual Problem

$$f^* = \inf_{x \in \mathbb{R}^n} \quad f(x)$$

Primal

$$\text{s.t.} \quad h_i(x) \geq 0, \; i = 1, \ldots, m$$

$$\ell_j(x) = 0, \; j = 1, \ldots, r$$

Since dual objective $g(\alpha, \beta)$ gives a lower bound, the best lower bound can be obtained by maximizing it for all dual feasible variables:

Dual

$$g^* = \sup_{\alpha \in \mathbb{R}^m, \beta \in \mathbb{R}^r} \quad g(\alpha, \beta)$$

$$\text{s.t.} \quad \alpha \geq 0$$

Weak duality: $f^* \geq g^*$ Always true!

Dual Problem is Always Convex Optimization

$$g(\alpha, \beta) = \min_{x \in \mathbb{R}^n} \left\{ f(x) - \sum_{i=1}^{m} \alpha_i h_i(x) - \sum_{j=1}^{r} \beta_j \ell_j(x) \right\}$$

$$= -\max_{x \in \mathbb{R}^n} \left\{ -f(x) + \sum_{i=1}^{m} \alpha_i h_i(x) + \sum_{j=1}^{r} \beta_j \ell_j(x) \right\}$$

Pointwise maximum of convex (affine) functions in (α, β)

Therefore g is concave in (α, β). With $\alpha \geq 0$ (convex constraint), this implies that the dual problem is always a convex optimization, even if the primal is not.

Duality Gap

For primal optimal x^* and dual optimal (α^*, β^*),

$$\text{(duality gap)} := f(x^*) - g(\alpha^*, \beta^*) \geq 0$$

Strong Duality

$$\text{Strong Duality} \;\Leftrightarrow\; f^* = g^* \;\Leftrightarrow\; \text{No Duality Gap}$$

For convex primal problems, we have strong duality if

Slater's condition holds: there exists at least one strictly feasible point in the primal

Strong duality may hold in nonconvex problems

$$\min_{x} \quad -x^T x$$

$$\text{s.t.} \;\; x^T x \leq 1$$

$$\mathcal{L}(x,\lambda) = -x^T x - \lambda(1 - x^T x) = (\lambda - 1)x^T x - \lambda$$

$$q(\lambda) = \begin{cases} -\lambda & \lambda \geq 1 \\ -\infty & \text{o.w.} \end{cases}$$

Dual problem: $\max_{\lambda \geq 1} \; -\lambda$

There's no duality gap!!

Constraint Qualification (CQ)

CQ is required so that <u>Lagrange multipliers will exist</u> satisfying the KKT conditions

- **LICQ (Linear independence CQ)**: the gradients of active constraints are linearly independent at x*
  - → Lagrange multipliers exist and are **unique**

- **MFCQ (Mangasarian-Fromovitz CQ):**

  there exists $w \in \mathbb{R}^n$ s.t.

$$\nabla h_i(x^*)^T w > 0, \quad \text{for all active inequality constraints}$$
$$\nabla \ell_j(x^*)^T w = 0, \quad \text{for all equality constraints,}$$

  and the set of equality constraint gradients is linearly independent.

Slater's Condition

$$\inf_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad h_i(x) \geq 0,\ i = 1, \ldots, m$$
$$\ell_j(x) = 0,\ j = 1, \ldots, r$$

$$\mathcal{D} = \text{dom}\, f \cap \bigcap_{i=1}^{m} \text{dom}\, h_i \cap \bigcap_{j=1}^{r} \text{dom}\, \ell_j \neq \emptyset$$

- **Slater's condition:**

  there exists $x \in \text{relint}\, \mathcal{D}$ s.t.

$$\begin{cases} h_i(x) & > 0, \quad \text{for all (non-affine) inequality constraints} \\ \ell_j(x) & = 0, \quad \text{for all equality constraints.} \end{cases}$$

Convex opt & Slater's condition $\Rightarrow$ strong duality

Ex. Convex Opt Alone Is Not Enough

$$p^* = \min_{x,y>0} e^{-x} \text{ s.t. } x^2/y \le 0 \qquad \mathcal{D} = \{(x,y) \in \mathbb{R}^2 : y > 0\}$$

$$\mathcal{L}(x,y,\lambda) = e^{-x} + \lambda x^2/y$$

$$g(\lambda) = \inf_{(x,y)\in\mathcal{D}} (e^{-x} + \lambda x^2/y) = \begin{cases} 0 & \text{if } \lambda \ge 0 \\ -\infty & \text{if } \lambda < 0 \end{cases}$$

$$d^* = \max_{\lambda \ge 0} 0$$

$$p^* - d^* = 1 - 0 = 1$$

Karush-Kuhn-Tucker (KKT) Conditions

$$\min_{x \in \mathbb{R}^n} \; f(x)$$

$$\begin{aligned} 0 \le u_i \sim \quad & \text{s.t.} \;\; h_i(x) \ge 0, \; i = 1, \ldots, m \\ v_i \sim \quad & \qquad \ell_j(x) = 0, \; j = 1, \ldots, r \end{aligned}$$

C : feasible set

$(\tilde{x}; \tilde{u}, \tilde{v})$ satisfies the KKT if all of the following conditions are true:

$$0 = \nabla_x \mathcal{L}(\tilde{x}; \tilde{u}, \tilde{v}) = \nabla f(x^*) - \sum_{i=1}^{m} \tilde{u}_i \nabla h_i(x^*) - \sum_{j=1}^{r} \tilde{v}_j \nabla \ell_j(\tilde{x})$$ Lagrange optimality

$$h_i(\tilde{x}) \ge 0, \; \ell_j(\tilde{x}) = 0 \;\; \forall i, j$$ Primal feasibility

$$\tilde{u}_i \ge 0, \;\; \forall i$$ Dual feasibility

$$\tilde{u}_i \, h_i(\tilde{x}) = 0, \;\; \forall i$$ Complementary slackness

Optimality / Duality

Considerations:

1. When do optimal Lagrange multipliers exist ?
2. What is the relation between

x^* primal optimal (u^*, v^*) dual optimal	$\sim$	$(\tilde{x}, \tilde{u}, \tilde{v})$ satisfying the KKT conditions

3. When can we solve a dual instead of its primal, and obtain primal solutions from the dual solutions?

First-Order Necessary Optimality Condition (FONC)

Let x^* be a (local) minimizer, at which CQ holds. Then there exists Lagrange multipliers $(\tilde{u}, \tilde{v})$ satisfying the KKT conditions at $(x^*, \tilde{u}, \tilde{v})$.

x^* primal optimal

CQ →

↓

Multiplier $(\tilde{u}, \tilde{v})$ exists for which $(x^*, \tilde{u}, \tilde{v})$ sat KKT

FOC (Under Strong Duality)

x^* primal optimal

(u^*, v^*) dual optimal

Strong duality

Primal is convex opt

(x^*, u^*, v^*) satisfies KKT

FONC (Under Strong Duality)

Let x^* and (u^*, v^*) be primal and dual solutions satisfying strong duality. Then (x^*, u^*, v^*) satisfies the KKT conditions.

First, x^* and (u^*, v^*) are primal and dual feasible.

$$
\begin{aligned}
f(x^*) &= g(u^*, v^*) \\
&= \min_{x \in \mathbb{R}^n} \mathcal{L}(x; u^*, v^*) \\
&\leq \mathcal{L}(x^*; u^*, v^*) \\
&\leq f(x^*)
\end{aligned}
$$

Therefore, all inequalities should hold as equalities.

FONC (Under Strong Duality)

$$\begin{aligned} f(x^*) &= g(u^*, v^*) \\ &= \min_{x \in \mathbb{R}^n} \mathcal{L}(x; u^*, v^*) \\ &= \mathcal{L}(x^*; u^*, v^*) \\ &= f(x^*) \end{aligned}$$

x^* minimizes $\mathcal{L}(x; u^*, v^*)$, and thus is a stationary point. i.e. $0 \in \partial_x \mathcal{L}(x^*; u^*, v^*)$

$$\mathcal{L}(x^*; u^*, v^*) = f(x^*) - \sum_{i=1}^{m} \underbrace{u_i^*}_{\geq 0} \underbrace{h_i(x^*)}_{\geq 0} - \sum_{j=1}^{n} v_j^* \underbrace{\ell_j(x^*)}_{=0}$$

$$\Rightarrow \quad u_i^* h_i(x^*) = 0 \text{ should hold for all } i.$$

(No assumption on the convexity of the problem!)

Sufficient Optimality Condition (Primal is Convex Opt)

Let $\tilde{x}$ and $(\tilde{u}, \tilde{v})$ satisfy the KKT conditions.

Then, the duality gap is zero: $\tilde{x}$ and $(\tilde{u}, \tilde{v})$ are primal and dual solutions.

$$g(\tilde{u}, \tilde{v}) = \min_{x \in \mathbb{R}^n} \mathcal{L}(x; \tilde{u}, \tilde{v})$$

$$= f(\tilde{x}) - \sum_{i=1}^{m} \underbrace{\tilde{u}_i h_i(\tilde{x})}_{=0 \text{ (CS)}} - \sum_{j=1}^{r} \tilde{v}_j \underbrace{\ell_j(\tilde{x})}_{=0 \text{ (feasibility)}}$$

$$= f(\tilde{x})$$

The primal is a convex opt: $f(x)$ convex, $h_i(x)$ concave, $\ell_j(x)$ affine

$\Rightarrow \mathcal{L}(x; \tilde{u}, \tilde{v})$ is convex in x

$\Rightarrow 0 \in \partial_x \mathcal{L}(\tilde{x}; \tilde{u}, \tilde{v})$ is sufficient $\tilde{x}$ to be a minimizer of $\mathcal{L}(x; \tilde{u}, \tilde{v})$

Sufficient Optimality Condition (Primal is Convex Opt)

Let $\tilde{x}$ and $(\tilde{u}, \tilde{v})$ satisfy the KKT conditions.

Then, the duality gap is zero: $\tilde{x}$ and $(\tilde{u}, \tilde{v})$ are primal and dual solutions.

Q: why $\tilde{x}$ primal optimal?

$$
\begin{aligned}
f(\tilde{x}) &= \min_{x \in \mathbb{R}^n} \mathcal{L}(x; \tilde{u}, \tilde{v}) = \min_{x \in \mathbb{R}^n} \{ f(x) - \tilde{u}^T h(x) - \tilde{v}^T \ell(x) \} \\
&\leq \min_{x \in \mathcal{C}} \{ f(x) - \tilde{u}^T h(x) - \tilde{v}^T \ell(x) \} \\
&\leq \min_{x \in \mathcal{C}} f(x)
\end{aligned}
$$

FOC (Under Strong Duality)

x^* primal optimal

(u^*, v^*) dual optimal

Strong duality

(primal is convex opt)

(x^*, u^*, v^*) satisfies KKT

Strong Duality: Dual → Primal

An implication of the proof of "FONC + strong duality": given a dual solution (u*,v*), a primal solution x* is also a solution of

$$\min_{x \in \mathbb{R}^n} \mathcal{L}(x; u^*, v^*)$$

If $\mathcal{L}(x; u^*, v^*)$ is convex in x, then x* can be found by solving

$$0 \in \partial_x \mathcal{L}(x^*; u^*, v^*)$$

If $\min_{x \in \mathbb{R}^n} \mathcal{L}(x; u^*, v^*)$ has a unique solution, then it must be the unique primal solution

Fenchel Conjugate

$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, not necessarily convex

$f \not\equiv +\infty$, there exists an affine function minorizing f on $\mathbb{R}^n$

$\Rightarrow f(x) > -\infty \quad \forall x \qquad \text{dom}\, f := \{x : f(x) < +\infty\} \neq \emptyset$

$f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is the **conjugate** of f defined by

$$f^*(y) := \sup_{x \in \text{dom}\, f} \{y^T x - f(x)\}$$

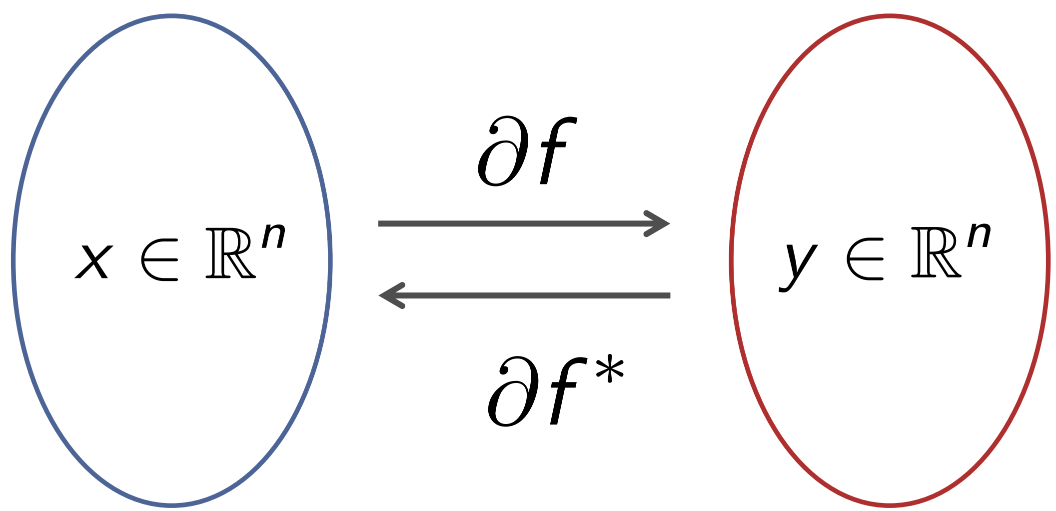
The mapping $f \mapsto f^*$ is called the *conjugacy* operation, *conjugation*, or *Legendre-Fenchel* transform.

$f^*(y)$ is always closed and convex

Why Conjugate ?

Space of Points
(Primal space)

Space of Gradients
(Dual space)



$$f^*(y) := \sup_{x \in \text{dom } f} \{y^T x - f(x)\}$$

(under some technical conditions, to be discussed)

Calculus Rules I

$$\alpha \in \mathbb{R}$$

$$g(x) = f(x) + \alpha \qquad \Rightarrow g^*(y) = f^*(y) - \alpha$$

$$g(x) = f(x - x_0) \qquad \Rightarrow g^*(y) = f^*(y) + y^T x_0$$

$$g(x) = f(x) + y_0^T x \qquad \Rightarrow g^*(y) = f^*(y - y_0)$$

$$g(x) = f(x) + y_0^T x + \alpha$$

Calculus Rules I

$$\alpha \in \mathbb{R}$$

$$g(x) = f(\alpha x),\ \alpha \neq 0 \quad \Rightarrow g^*(y) = f^*(y/\alpha)$$

$$g(x) = \alpha f(x),\ \alpha > 0 \quad \Rightarrow g^*(y) = \alpha f^*(y/\alpha)$$

$$g(x) = \alpha f(x/\alpha),\ \alpha > 0 \quad \Rightarrow g^*(y) = \alpha f^*(y)$$

Calculus Rules III

Separable sum:

$$f(x_1, x_2) = g(x_1) + h(x_2) \quad \Rightarrow \quad f^*(y_1, y_2) = g^*(y_1) + h^*(y_2)$$

Linear composition (*A* invertible):

$$g(x) = f(Ax) \quad \Rightarrow \quad g^*(y) = f^*(A^{-T} y)$$

Infimal convolution:

$$f(x) = \inf_{u+v=x} (g(u) + h(v)) \quad \Rightarrow \quad f^*(y) = g^*(y) + h^*(y)$$

Convexity

$\mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2 \neq \emptyset, \ \alpha \in [0,1],$

$$[\alpha f_1 + (1-\alpha) f_2]^* \leq \alpha f_1^* + (1-\alpha) f_2^*$$

Fenchel-Young Inequality

$\forall (x, y) \in \mathrm{dom}\, f \times \mathbb{R}^n,$

$$f(x) + f^*(y) \geq x^T y.$$

Equality holds if *y* is a subgradient of *f* at *x*, $y \in \partial f(x)$

Inequality: obvious from the definition.

If $y \in \partial f(x), f(x') - f(x) \geq y^T (x' - x) \;\; \forall x'.$

Therefore $y^T x - f(x) \geq \sup_{x'} \{ y^T x' - f(x') \} = f^*(y)$

Ex. Exponentiation

$$f(x) = \exp(x)$$

$$f^*(y) = \begin{cases} -\infty & y < 0 \\ 0 & y = 0 \\ y\log(y) - y & y > 0 \end{cases}$$

Ex. Negative Entropy

$$f(x) = \sum_{i=1}^{n} x_i \log(x_i)$$

$$f^*(y) = \sum_{i=1}^{n} e^{y_i - 1}$$

Ex. Indicator Function

$$f(x) = I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{o.w.} \end{cases}$$

Conjugate:

$$f^*(y) = I_C^*(y) = \sup_{x \in C} \; y^T x$$

This f^* is called as the **support function** of the set C

Ex. Norms

$$f(x) = \|x\| \qquad\qquad f^*(y) = I_{\|\cdot\|_* \le 1}(y)$$

where $\|y\|_* := \max_{\|z\| \le 1} z^T y$ is the **dual norm** of $\|\cdot\|$

$$f^*(y) := \sup_{x \in \mathbb{R}^n} \{x^T y - \|x\|\}$$

(Hölder's ineq.)

if $\|y\|_* \le 1$, then $x^T y - \|x\| \le \|x\| \|y\|_* - \|x\| \le 0$

if $\|y\|_* > 1$, consider $\tilde{z} \in \mathbb{R}^n : \|\tilde{z}\| \le 1$ and $\tilde{z}^T y = \|y\|_*$,

$$(t\tilde{z})^T y - \|t\tilde{z}\| = t(\tilde{z}^T y - \|\tilde{z}\|) \to \infty \;\text{ with } t \to \infty$$

Biconjugation

$$f^{**}(x) = (f^*)^*(x) = \sup_{y \in \mathbb{R}^n} \{x^T y - f^*(y)\}$$

$$f^{**} \leq f \quad \Leftrightarrow \quad \text{epi } f^{**} \supseteq \text{epi } f$$

$$\text{epi } f^{**} = \text{cl conv epi } f$$

If f is convex and closed, $f^{**} = f$

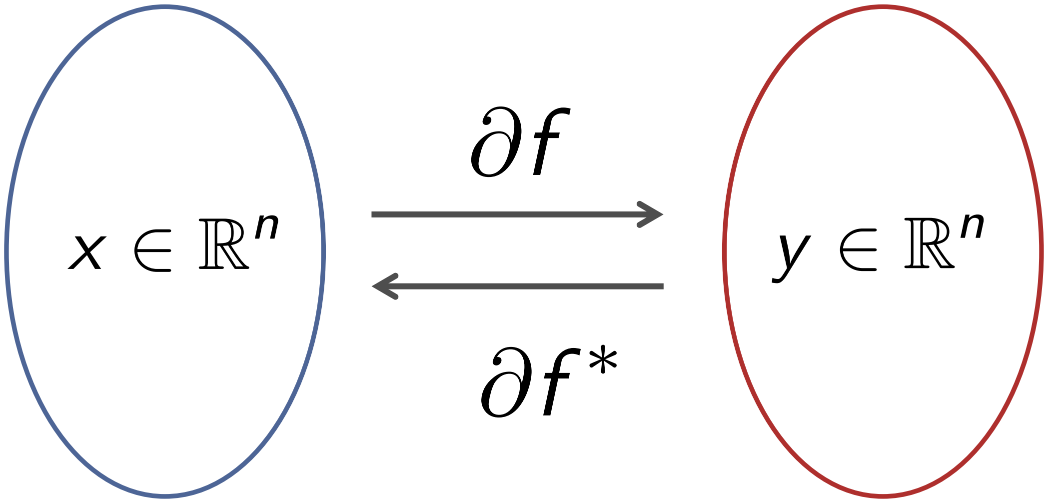
Note that f^* always satisfies the required conditions for conjugation.

Subgradient Connection

If f is convex and closed,

$$y \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(y)$$

$$\text{also,} \quad \Leftrightarrow f(x) + f^*(y) = x^T y$$



Strong Convexity & Smoothness: Duality

f closed and strongly convex with a constant $\alpha > 0$:

$$\text{dom}\, f^* = \mathbb{R}^n$$

$$\nabla f^*(y) = \underset{x \in \text{dom}\, f}{\arg\max}\, \{y^T x - f(x)\}, \quad \forall y \in \mathbb{R}^n$$

$\nabla f^*(y)$ is Lipschitz continuous with the constant $1/\alpha$

This gives the fundamental idea of so-called "Nesterov's smoothing"

Ex. Dual of Lasso

$A \in \mathbb{R}^{m \times n}$

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1$$

$$\min_{x \in \mathbb{R}^n, z \in \mathbb{R}^m} \frac{1}{2}\|y - z\|_2^2 + \lambda\|x\|_1, \text{ s.t. } z = Ax$$

Dual objective:

$$g(u) = \inf_{x \in \mathbb{R}^n, z \in \mathbb{R}^m} \frac{1}{2}\|y - z\|^2 + \lambda\|x\|_1 + u^T(z - Ax)$$

$$= \inf_{z \in \mathbb{R}^m} \left\{ \frac{1}{2}\|y - z\|^2 + u^T z \right\} + \inf_{x \in \mathbb{R}^n} \{\lambda\|x\|_1 - u^T Ax\}$$

$$= \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y - u\|^2 - \lambda \sup_{x \in \mathbb{R}^n} \{v^T x - \|x\|_1\}, \quad v := A^T u/\lambda$$

$$g(u) = \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y-u\|^2 - \lambda \underbrace{\sup_{x\in\mathbb{R}^n} \{v^T x - \|x\|_1\}}, \quad v := A^T u/\lambda$$

Conjugate of $\|x\|_1$

$$f(x) = \|x\| \qquad f^*(v) = \begin{cases} 0 & \text{if } \|v\|_* \le 1 \\ \infty & \text{o.w.} \end{cases}$$

Dual norm: $$\|v\|_* = \sup_{x:\|x\|\le 1} v^T x = \sup_{x:x\neq 0} \frac{v^T x}{\|x\|}$$

$$\|\cdot\|_1 \longleftrightarrow \|\cdot\|_\infty$$
$$\|\cdot\|_2 \longleftrightarrow \|\cdot\|_2$$
$$\|\cdot\|_p \longleftrightarrow \|\cdot\|_q, \quad p,q \ge 1, 1/p + 1/q = 1.$$

Dual problem:

$$\sup_{u\in\mathbb{R}^n} g(u) = \sup_{u\in\mathbb{R}^n} \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y-u\|^2 - \lambda \begin{cases} 0 & \|v\|_\infty \le 1 \\ +\infty & \text{o.w.} \end{cases}, \quad v := A^T u/\lambda$$

$$\max_{u\in\mathbb{R}^n} -\frac{1}{2}\|y-u\|^2 \ \text{ s.t. } \ \|A^T u\|_\infty \le \lambda$$

Or, equivalently, $-\min_{u\in\mathbb{R}^n} \frac{1}{2}\|y-u\|^2 \ \text{ s.t. } \ \|A^T u\|_\infty \le \lambda$

How to solve this?

Convex opt + Slater's condition ➔ strong duality holds

Given a dual solution u^*, we can find a primal soln by solving

$$\nabla_z \mathcal{L}(x, z^*; u^*) = -(y - z^*) + u^* = 0$$

$\Rightarrow z^* = y - u^* \quad \Rightarrow \quad$ Solve for x^*: $Ax^* = z^*$.

Ex. Fused Lasso

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2}\|b - x\|_2^2 + \lambda \|Dx\|_1$$

$D \in \mathbb{R}^{m \times n}$: a penalty matrix

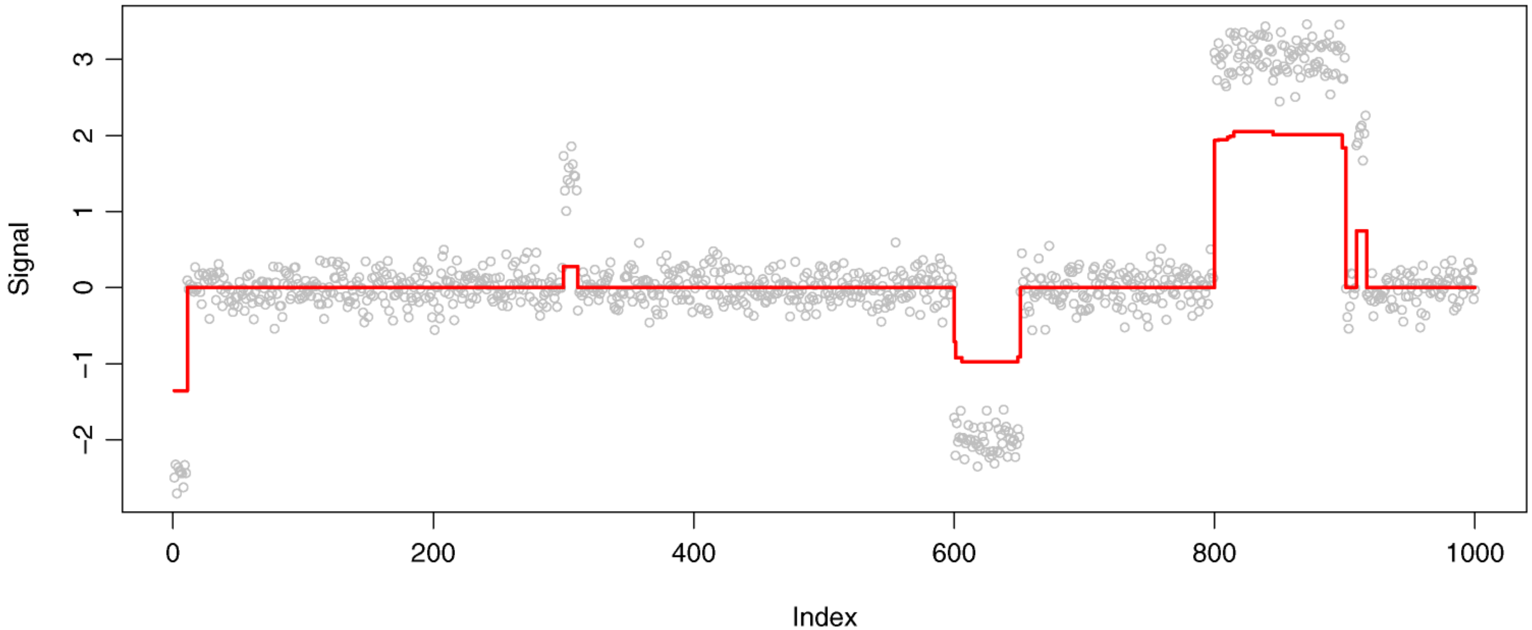
Ex. 1-D Fused Lasso

$$D = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix}$$

Ex. D is an incident matrix for a graph $G = (\{1, \ldots, n\}, E)$,

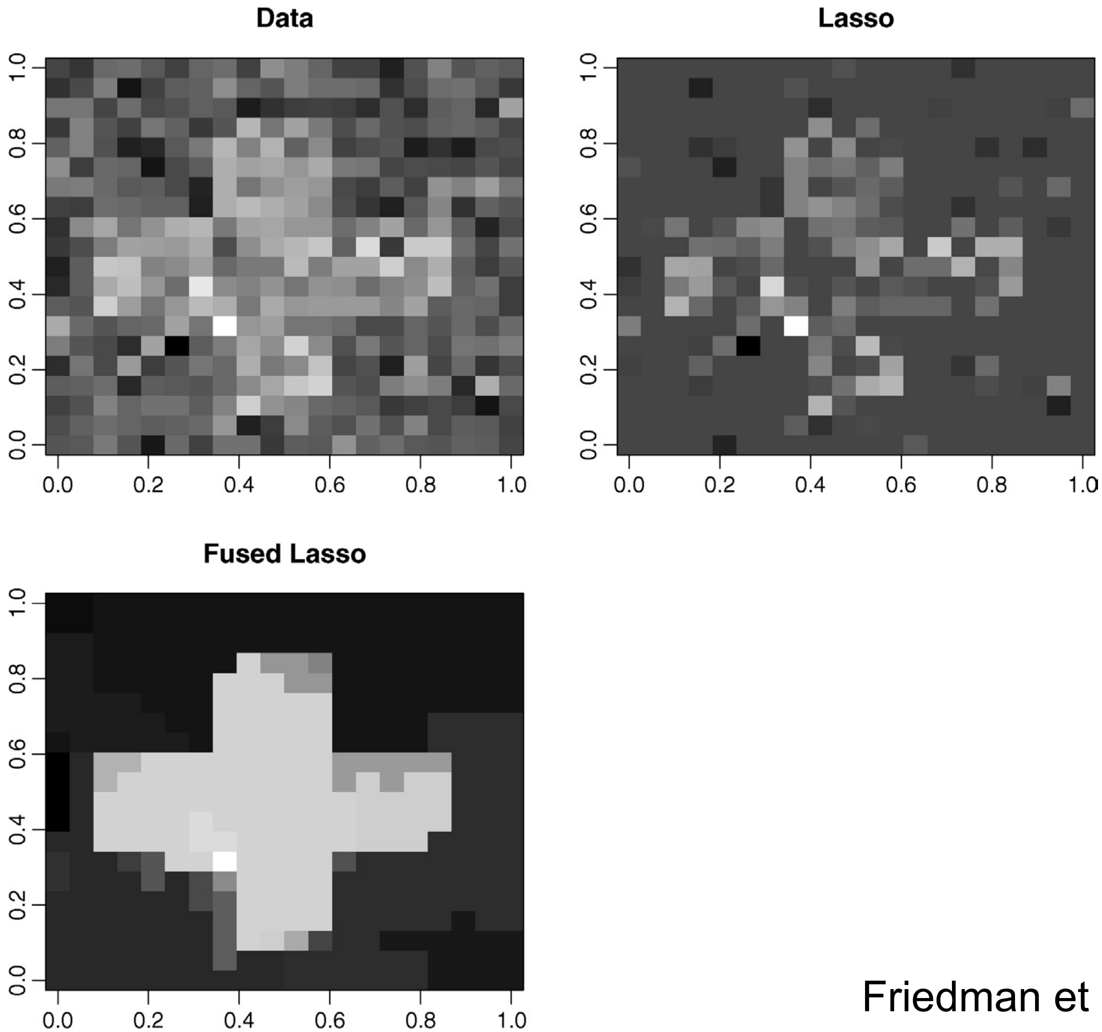
$$\|Dx\|_1 = \sum_{(i,j) \in E} |x_i - x_j|$$

1-D Fused Lasso



Friedman et al., *Ann. Appl. Stat.,* 2007

2-D Fused Lasso



Friedman et al., *Ann. Appl. Stat.*, 2007

Fused Lasso: Dual

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2}\|b - x\|_2^2 + \lambda \|Dx\|_1 \qquad\qquad D \in \mathbb{R}^{m \times n} : \text{a penalty matrix}$$

FISTA? The regularization term is not separable in general, so prox operation may not be simple

ADMM approach is possible, e.g. using z = Dx

$$\min_{x \in \mathbb{R}^n, z \in \mathbb{R}^m} \quad \frac{1}{2}\|b - x\|_2^2 + \lambda \|z\|_1$$

$$\text{s.t.} \quad z = Dx$$

Or, we can consider the dual problem (homework)