A Perturbation View of Level-Set Methods

Things get weird without strong duality

Ron Estrin
ICME, Stanford University

SIAM CSE 2019
Feb. 26, 2019
Joint Work with Michael Friedlander

The Problem

Want to solve:

$$\tau_p^* = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \leq 0 \}$$

with f, c convex

- $f(x)$ is 'simple'
- $c(x)$ is 'complicated'

The Problem

Want to solve:

$$\tau_p^* = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \leq 0 \}$$

with f, c convex

- $f(x)$ is ‘simple’
- $c(x)$ is ‘complicated’

Example (**sparse optimization**):

$$\min_x \|x\|_1 \quad \text{subject to} \quad \|Ax - b\|_2^2 - \sigma^2 \leq 0$$

Flip it on its head!

$$\tau_p^* = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \le 0 \} \qquad \left(\min_x \|x\|_1 \text{ s.t. } \|Ax - b\|_2^2 - \sigma^2 \le 0 \right)$$

$$\Downarrow$$

$$v(\tau) = \inf_{x \in \mathcal{X}} \{ c(x) \mid f(x) \le \tau \} \qquad \left(\min_x \|Ax - b\|_2^2 - \sigma^2 \text{ s.t. } \|x\|_1 \le \tau \right)$$

Flipped problem $v(\tau)$ is 'easier' to solve (for fixed τ)

e.g., using projected gradient descent

Flip it on its head!

$$\tau_p^* = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \leq 0 \} \qquad \left(\min_x \|x\|_1 \text{ s.t. } \|Ax - b\|_2^2 - \sigma^2 \leq 0 \right)$$

$$\Downarrow$$

$$v(\tau) = \inf_{x \in \mathcal{X}} \{ c(x) \mid f(x) \leq \tau \} \qquad \left(\min_x \|Ax - b\|_2^2 - \sigma^2 \text{ s.t. } \|x\|_1 \leq \tau \right)$$

Flipped problem $v(\tau)$ is 'easier' to solve (for fixed τ)

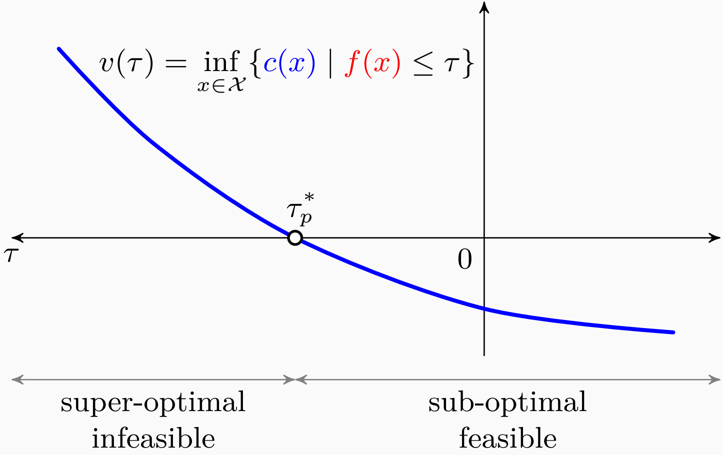
e.g., using projected gradient descent

Interpretation:

How much do I have to violate the constraints to get a certain objective value?

The Level-Set Method

Find τ such that $v(\tau) = 0$.



Developed by (van den Berg and Friedlander, 2008;

Aravkin, Burke, Drusvyatskiy, Friedlander, Roy, 2016).

A weird problem

3×3 SDP:

$$\min_{X \succeq 0} \; -2X_{31} \;\text{ subject to }\; X_{11} = 0,\; X_{22} + 2X_{31} = 1;\;\; X = \begin{bmatrix} x_{11} & \cdot & \cdot \\ x_{21} & x_{22} & \cdot \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

A weird problem

3×3 SDP:

$$\min_{X \succeq 0} \; -2X_{31} \;\text{ subject to }\; X_{11} = 0,\; X_{22} + 2X_{31} = 1; \;\; X = \begin{bmatrix} X_{11} & \cdot & \cdot \\ X_{21} & X_{22} & \cdot \\ X_{31} & X_{32} & X_{33} \end{bmatrix}$$

Solution $X^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, optimal value $\tau_p^* = 0$.

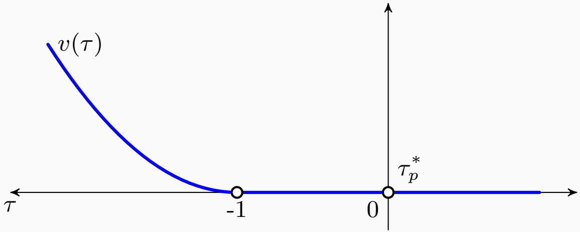
A weird problem

3×3 SDP:

$$\min_{X \succeq 0} \; -2X_{31} \;\text{ subject to }\; X_{11} = 0,\; X_{22} + 2X_{31} = 1; \;\; X = \begin{bmatrix} X_{11} & \cdot & \cdot \\ X_{21} & X_{22} & \cdot \\ X_{31} & X_{32} & X_{33} \end{bmatrix}$$

Solution $X^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, optimal value $\tau_p^* = 0$.

But $v(\tau) = \inf_{X \succeq 0} \left\{ |X_{11}|^2 + |X_{22} + 2X_{31} - 1|^2 \mid -2X_{31} \le \tau \right\}$ looks like



We get the wrong answer! Why?

Digging a little deeper...

$$v(\tau) = \inf_{X \succeq 0} \left\{ |X_{11}|^2 + |X_{22} + 2X_{31} - 1|^2 \mid -2X_{31} \le \tau \right\}$$

Suppose $-1 < \tau < 0$:

- Let $\epsilon > 0$, define $X(\epsilon) = \begin{bmatrix} \epsilon & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4\epsilon^2} \end{bmatrix}$

Digging a little deeper...

$$v(\tau) = \inf_{X \succeq 0} \left\{ |X_{11}|^2 + |X_{22} + 2X_{31} - 1|^2 \mid -2X_{31} \leq \tau \right\}$$

Suppose $-1 < \tau < 0$:

- Let $\epsilon > 0$, define $X(\epsilon) = \begin{bmatrix} \epsilon & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4\epsilon^2} \end{bmatrix}$
- Then $f(X(\epsilon)) \leq \tau$ and $c(X(\epsilon)) = \epsilon^2$

Digging a little deeper...

$$v(\tau) = \inf_{X \succeq 0} \left\{ |X_{11}|^2 + |X_{22} + 2X_{31} - 1|^2 \mid -2X_{31} \leq \tau \right\}$$

Suppose $-1 < \tau < 0$:

- Let $\epsilon > 0$, define $X(\epsilon) = \begin{bmatrix} \epsilon & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4\epsilon^2} \end{bmatrix}$
- Then $f(X(\epsilon)) \leq \tau$ and $c(X(\epsilon)) = \epsilon^2$
- As $\epsilon \to 0$, $c(X(\epsilon)) \to 0$, but $X(0)$ is **not feasible**!

Even though $v(\tau) = 0$, there does not exist a feasible point s.t. $f(x) \leq \tau$!

Duality

Common description of duality in courses on convex analysis:

1. Lagrangian: $L(x, y) = f(x) + c(x)^T y$
2. Dual function: $g(y) = \min_{x \in \mathcal{X}} L(x, y)$
3. Dual problem:

$$\tau_d^* = \sup_{y \geq 0} g(y)$$

Duality

Common description of duality in courses on convex analysis:

1. Lagrangian: $L(x, y) = f(x) + c(x)^T y$
2. Dual function: $g(y) = \min_{x \in \mathcal{X}} L(x, y)$
3. Dual problem:

$$\tau_d^* = \sup_{y \geq 0} g(y)$$

Weak duality: $\tau_d^* \leq \tau_p^*$ (always holds)
Strong duality: $\tau_d^* = \tau_p^*$ (sometimes holds)

Duality

Common description of duality in courses on convex analysis:

1. Lagrangian: $L(x, y) = f(x) + c(x)^T y$
2. Dual function: $g(y) = \min_{x \in \mathcal{X}} L(x, y)$
3. Dual problem:

$$\tau_d^* = \sup_{y \geq 0} g(y)$$

Weak duality: $\tau_d^* \leq \tau_p^*$ (always holds)
Strong duality: $\tau_d^* = \tau_p^*$ (sometimes holds)

For the 3×3 SDP example:

$$\tau_d^* = -1$$

Coincidence?!

A geometric view of duality (Veinott 1989)

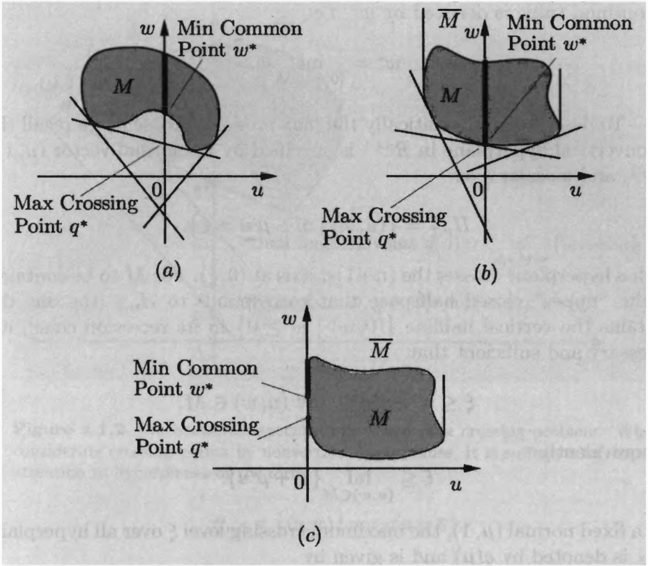


Figure 4.1.1, Convex Optimization Theory, Bersekas (2009)

A geometric view of duality (Veinott 1989)

Perturbed problem:

$$p(u) = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \le u \},$$

Let $M = \text{epi} p = \{(u, \alpha) \mid p(u) \le \alpha\}$.

$$\text{Min Common Point: } p(0) = \tau_p^*$$

$$\text{Max Crossing Point: } \tau_d^*$$

A geometric view of duality (Veinott 1989)

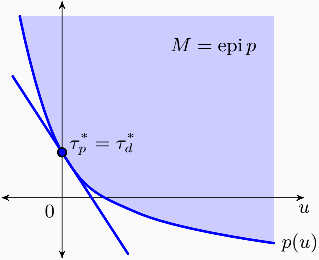
Perturbed problem:

$$p(u) = \inf_{x \in \mathcal{X}} \{ f(x) \mid c(x) \le u \},$$

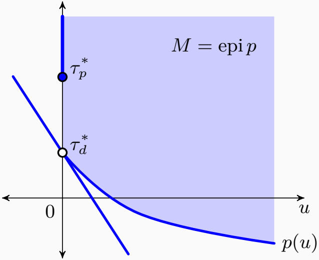
Let $M = \text{epi} p = \{(u, \alpha) \mid p(u) \le \alpha\}$.

Min Common Point: $p(0) = \tau_p^*$

Max Crossing Point: τ_d^*



(a) Strong Duality

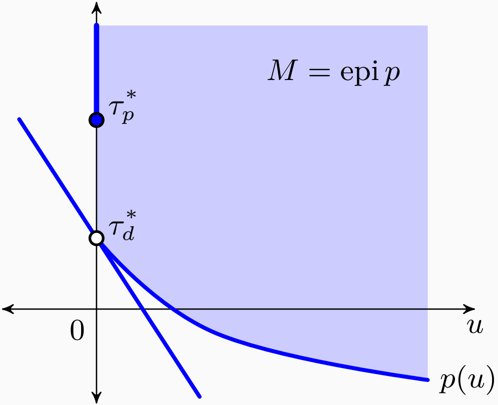


(b) No Strong Duality

So if strong duality fails...

For $\tau_d^* < \tau < \tau_p^*$:

- For any $u > 0$, $p(u) \leq \tau$
- Implies that $\exists x$ such that $c(x) \leq u$ and $f(x) \leq \tau$
- Thus $v(\tau) \leq u$ for all $u > 0 \implies v(\tau) = 0$



The Main Result

Theorem

$$\tau > \tau_d^* \implies v(\tau) = \inf_{x \in \mathcal{X}} \{c(x) \mid f(x) \leq \tau\} = 0$$

The Main Result

Theorem

$$\tau > \tau_d^* \implies v(\tau) = \inf_{x \in \mathcal{X}} \{c(x) \mid f(x) \leq \tau\} = 0$$

Corollary

If strong duality fails ($\tau_d^* < \tau_p^*$), the level-set method fails.

The Main Result

Theorem

$$\tau > \tau_d^* \implies v(\tau) = \inf_{x \in \mathcal{X}} \{c(x) \mid f(x) \le \tau\} = 0$$

Corollary

If strong duality fails ($\tau_d^* < \tau_p^*$), the level-set method fails.

Corollary

If $f(x)$ is coercive ($f(x) \to \infty$ as $\|x\| \to \infty$), or $\mathcal{X}$ is compact, then **strong duality holds**.

Conclusion

Open Question:

We know that the level-set method fails without strong duality (as many methods do), how can we fix this?

Possible idea: proximal-point + level-set method

References:

- E. van den Berg and M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. November 2016
- A. Y. Aravkin, J. V. Burke, D. Drusvyatskiy, M. P. Friedlander, S. Roy. Level-set methods for convex optimization. February 2016
- R. Estrin and M. P. Friedlander. A perturbation view of level-set methods for convex optimization. Submitted to Math. Comp., 2018

Thank you for your attention!