On the Primal-Dual Geometry of Level Sets in Linear and Conic Optimization

Robert M. Freund
M.I.T.

October, 2009

Primal and Dual Linear Optimization Problems

\[ P : \text{VAL} := \min_{x} c^T x \quad D : \text{VAL} := \max_{y,z} b^T y \]

\[
\begin{align*}
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0 \\
\text{s.t.} & \quad A^T y + z = c \\
& \quad z \geq 0
\end{align*}
\]

\[ A \in \mathbb{R}^{m \times n} \]

“\( x \geq 0 \)” is “\( x \in \mathbb{R}^n_+ \)”
Meta-Lessons from Interior-Point Theory/Methods

• Linear optimization is not much more special than conic convex optimization

• The problem is ill-conditioned if VAL is finite but the primal or dual objective function level sets are unbounded

• $\varepsilon$-optimal solutions are important objects

• Choice of norm is important; some norms are more natural for certain settings
Meta-Lessons from Interior-Point Theory/Methods, continued

• All the important activity is in the (regular) cones

Indeed, we could eliminate the $y$-variable and re-write $P$ and $D$ as:

$$
P : \min_x \quad c^T x \quad D : \text{VAL} := \min_z \quad (x^0)^T z$$

such that:

- $x - x^0 \in L$
- $x \geq 0$

- $z - c \in L^\perp$
- $z \geq 0$

where $x^0$ satisfies $Ax^0 = b$ and $L = \text{null}(A)$.

But we won’t.
Primal and Dual Level Sets

\[ P : \text{VAL} := \min_x c^T x \quad \quad D : \text{VAL} := \max_{y,z} b^T y \]

\[
\begin{align*}
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0 \\
\text{s.t.} & \quad A^T y + z = c \\
& \quad z \geq 0
\end{align*}
\]

\[ P_\varepsilon := \{ x : Ax = b, x \geq 0, c^T x \leq \text{VAL} + \varepsilon \} \]

\[ D_\delta := \{ z : \exists y \text{ satisfying } A^T y + z = c, z \geq 0, b^T y \geq \text{VAL} - \delta \} \]
Level Set Geometry Measures

Define for $\varepsilon, \delta > 0$:

\[
R^P_\varepsilon := \max_x \|x\|_1 \quad \quad \quad \quad \quad \quad r^D_\delta := \max_{y,z} \min_j \{z_j\}
\]

s.t. $Ax = b$
$x \geq 0$
$c^T x \leq \text{VAL} + \varepsilon$

s.t. $A^T y + z = c$
$z \geq 0$
$b^T y \geq \text{VAL} - \delta$

$R^P_\varepsilon$ is the norm of the largest primal $\varepsilon$-optimal solution

$r^D_\delta$ measures the largest distance to the boundary of $\mathbb{R}^n_+$ among all dual $\delta$-optimal solutions $z$
Level Set Geometry Measures, continued

\[ R^{P}_{\varepsilon} := \max_{x} \|x\|_1 \quad r^{D}_{\delta} := \max_{z} \min_{j}\{z_j\} \]

\[ \text{s.t.} \quad x \in P_{\varepsilon} \quad \text{s.t.} \quad z \in D_{\delta} \]

\[ P_{\varepsilon} := \{x : Ax = b, x \geq 0, c^T x \leq \text{VAL} + \varepsilon\} \]

\[ D_{\delta} := \{z : \exists y \text{ satisfying } A^T y + z = c, z \geq 0, b^T y \geq \text{VAL} - \delta\} \]
$R_P^\varepsilon$ Measures Large Near-Optimal Solutions

$0^*$

c^T x = \text{VAL}_\varepsilon + \varepsilon$
$R^P_{\varepsilon}$ Measures Large Near-Optimal Solutions

\[ c^T x = \text{VAL}_* + \varepsilon \]

Feasible Region
$r^D_{\delta}$ Measures Nicely Interior Near-Optimal Solutions

\[ c^T x = \text{VAL}_\star + \varepsilon \]

\[ c^T x = \text{VAL}_\star + \varepsilon \]
Main Result: \( R_P^\varepsilon \) and \( r_D^\delta \) are Reciprocally Related

\[
R_P^\varepsilon := \max_x \|x\|_1 \quad \quad r_D^\delta := \max_z \min_j \{z_j\}
\]

s.t. \( x \in P_\varepsilon \) \quad \quad s.t. \( z \in D_\delta \)

Main Theorem: Suppose VAL is finite. If \( R_P^\varepsilon \) is positive and finite, then

\[
\min \{\varepsilon, \delta\} \leq R_P^\varepsilon \cdot r_D^\delta \leq \varepsilon + \delta.
\]

Otherwise \( \{R_P^\varepsilon, r_D^\delta\} = \{0, \infty\} \).
Comments

\[ \min\{\varepsilon, \delta\} \leq R^P_\varepsilon \cdot r^D_\delta \leq \varepsilon + \delta \]

- each inequality can be tight (and cannot be improved)

- setting \( \delta = \varepsilon \), we obtain \( \varepsilon \leq R^P_\varepsilon \cdot r^D_\varepsilon \leq 2\varepsilon \), showing these two measures are inversely proportional (to within a factor of 2)

- exchanging the roles of \( P \) and \( D \)

- how to prove
Comments, continued

\[ R^P_\varepsilon := \max_x \|x\|_1 \quad r^D_\delta := \max_z \min_j \{z_j\} \]

s.t. \( x \in P_\varepsilon \) \quad s.t. \( z \in D_\delta \)

\[ \varepsilon \leq R^P_\varepsilon \cdot r^D_\delta \leq 2\varepsilon \]

“The maximum norms of the primal objective level sets are almost exactly inversely proportional to the maximum distances to the boundary of the dual objective level sets”
Relation to LP Non-Regularity Property

**Standard LP Non-Regularity Property:** If VAL is finite, the set of primal optimal solutions is unbounded iff every dual feasible $z$ lies in the boundary of $\mathbb{R}_n^+$.  

$$R^P_\varepsilon := \max_x \|x\|_1 \quad r^D_\delta := \max_z \min_j \{z_j\}$$

s.t. $x \in P_\varepsilon$ \hspace{1cm} s.t. $z \in D_\delta$

In our notation, this is $R^P_\varepsilon = \infty$ iff $r^D_\delta = 0$ regardless of the values of $\varepsilon, \delta$ (which is the second part of the Main Theorem)
Relation to LP Non-Regularity, continued

\[ R_P^{\varepsilon} := \max_x \|x\|_1 \quad r_D^{\delta} := \max_z \min_j \{z_j\} \]

s.t. \( x \in P_\varepsilon \) \quad s.t. \( z \in D_\delta \)

The first part of the main theorem is: if \( R_P^{\varepsilon} \) is finite and positive, then

\[ \min\{\varepsilon, \delta\} \leq R_P^{\varepsilon} \cdot r_D^{\delta} \leq \varepsilon + \delta \]

This then is a generalization to nearly-non-regular problems, where \( R_P^{\varepsilon} \) is finite and \( r_D^{\delta} \) is non-zero
Question about Main Result

\[ R^P_\varepsilon := \max_x \|x\|_1 \quad r^D_\delta := \max_z \min_j \{z_j\} \]

s.t. \( x \in P_\varepsilon \) \quad s.t. \( z \in D_\delta \)

Q: Why the \( \| \cdot \|_1 \) norm?

A: Because \( f(x) := \|x\|_1 \) is a linear function on the cone \( \mathbb{R}^n_+ \). The linearity gives \( R^P_\varepsilon \) nice properties. If \( \| \cdot \| \) is not linear on \( \mathbb{R}^n_+ \) then we have to slightly weaken the main theorem as we will see . . . .
Primal and Dual Conic Problem

\[ P : \text{VAL}^* := \min_x c^T x \quad \quad \quad D : \text{VAL}^* := \max_{y,z} b^T y \]

\[ \text{s.t.} \quad Ax = b \quad \quad \quad \text{s.t.} \quad A^T y + z = c \]
\[ \quad x \in C \quad \quad \quad \quad \quad \quad z \in C^* \]

\( C \subset X \) is a regular cone: closed, convex, pointed, with nonempty interior

\( C^* := \{z : z^T x \geq 0 \ \forall x \in C\} \)
Primal and Dual Level Sets

\[ P : \text{VAL}_* := \min_x c^T x \quad D : \text{VAL}_* := \max_{y,z} b^T y \]

\[
\text{s.t.} \quad Ax = b \quad \text{s.t.} \quad A^T y + z = c \\
\quad x \in C \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad z \in C^* 0
\]

\[ P_\varepsilon := \{ x : Ax = b, x \in C, c^T x \leq \text{VAL}_* + \varepsilon \} \]

\[ D_\delta := \{ z : \exists y \text{ satisfying } A^T y + z = c, z \in C^*, b^T y \geq \text{VAL}_* - \delta \} \]
Level Set Geometry Measures

Define for $\varepsilon, \delta > 0$:

$$R^P_\varepsilon := \max_x \|x\|$$

s.t. $Ax = b$

$$x \in C$$

$$c^T x \leq \text{VAL}_* + \varepsilon$$

$$r^D_\delta := \max_{y,z} \text{dist}_*(z, \partial C^*)$$

s.t. $A^T y + z = c$

$$z \in C^*$$

$$b^T y \geq \text{VAL}_* - \delta$$

$\text{dist}_*(z, \partial C^*)$ denotes the distance from $z$ to $\partial C^*$ in the dual norm

$$\|z\|_* := \max\{z^T x : \|x\| \leq 1\}$$
Level Set Geometry Measures, continued

\[ R_{\varepsilon}^P := \max_x \|x\| \]
\[ r_\delta^D := \max_{y,z} \operatorname{dist}_*(z, \partial C^*) \]

s.t. \quad Ax = b \\
\quad x \in C \\
\quad c^T x \leq \text{VAL}_* + \varepsilon \\

s.t. \quad A^T y + z = c \\
\quad z \in C^* \\
\quad b^T y \geq \text{VAL}_* - \delta

\( R_{\varepsilon}^P \) is the norm of the largest primal \( \varepsilon \)-optimal solution

\( r_\delta^D \) measures the largest distance to the boundary of \( C^* \) among all dual \( \delta \)-optimal solutions \( z \)
Level Set Geometry Measures, continued

\[ R_{\varepsilon}^P := \max_x \|x\| \quad r_{\delta}^D := \max_z \text{dist}_*(z, \partial C^*) \]

s.t. \( x \in P_{\varepsilon} \) s.t. \( z \in D_{\delta} \)

\[ P_{\varepsilon} := \{ x : Ax = b, x \in C, c^T x \leq \text{VAL}_* + \varepsilon \} \]

\[ D_{\delta} := \{ z : \exists y \text{ satisfying } A^T y + z = c, z \in C^*, b^T y \geq \text{VAL}^* - \delta \} \]
$R^P_{\epsilon}$ Measures Large Near-Optimal Solutions

$c^T x = \text{VAL}_{\epsilon} + \epsilon$
$R^P_{\varepsilon}$ Measures Large Near-Optimal Solutions

$c^T x = \text{VAL}_* + \varepsilon$

Feasible Region
$r^D_\delta$ Measures Nicely Interior Near-Optimal Solutions

\[ c^T x = \text{VAL}_* + \varepsilon \]
Main Result, Again: $R_{\varepsilon}^P$ and $r_{\delta}^D$ are Reciprocally Related

$$R_{\varepsilon}^P := \max_x \|x\| \quad r_{\delta}^D := \max_z \text{dist}_*(z, \partial C^*)$$

s.t. $x \in P_{\varepsilon}$ \quad s.t. $z \in D_{\delta}$

Main Theorem: Suppose $\text{VAL}_*$ is finite. If $R_{\varepsilon}^P$ is positive and finite, then

$$\tau_{C^*} \cdot \min\{\varepsilon, \delta\} \leq R_{\varepsilon}^P \cdot r_{\delta}^D \leq \varepsilon + \delta.$$  

If $R_{\varepsilon}^P = 0$, then $r_{\delta}^D = \infty$. If $R_{\varepsilon}^P = \infty$ and $\text{VAL}_*$ is finite, then $r_{\delta}^D = 0$.

Here $\tau_{C^*}$ denotes the width of the cone $C^*$. . . . .
On the Width of a Cone

Let $K$ be a convex cone with nonempty interior

$$\tau_K := \max_x \{ \text{dist}(x, \partial K) : x \in K, \|x\| \leq 1 \}$$

If $K$ is a regular cone, then $\tau_K \in (0, 1]$

$\tau_K$ generalizes Goffin's "inner measure" for Euclidean norm
A Cone with small Width $\tau_K$

$\tau_K \ll 1$
Equivalence of Norm Linearity and Width of Polar Cone

**Proposition:** Let $K$ be a regular cone. The following statements are equivalent:

- $\tau_K^* \geq \alpha$, and
- there exists $\bar{w}$ for which
  \[ \alpha \bar{w}^T x \leq \|x\| \leq \bar{w}^T x \quad \text{for all } x \in K \]

**Corollary:** $\tau_K^* = 1$ implies $f(x) := \|x\|$ is linear on $K$
Main Result, Again: $R^P_{\varepsilon}$ and $r^D_\delta$ are Reciprocally Related

$$R^P_{\varepsilon} := \max_x \|x\| \quad \quad r^D_\delta := \max_z \text{dist}_*(z, \partial C^*)$$

s.t. $x \in P_{\varepsilon}$ \quad s.t. $z \in D_{\delta}$

Main Theorem: Suppose $\text{VAL}_{\ast}$ is finite. If $R^P_{\varepsilon}$ is positive and finite, then

$$\tau_{C^*} \cdot \min\{\varepsilon, \delta\} \leq R^P_{\varepsilon} \cdot r^D_\delta \leq \varepsilon + \delta .$$

If $R^P_{\varepsilon} = 0$, then $r^D_\delta = \infty$. If $R^P_{\varepsilon} = \infty$ and $\text{VAL}_{\ast}$ is finite, then $r^D_\delta = 0$.\[\square\]
Comments

$$\tau_{C^*} \cdot \min\{\varepsilon, \delta\} \leq R_{\varepsilon}^P \cdot r_D^\delta \leq \varepsilon + \delta$$

- each inequality can be tight (and cannot be improved)

- many naturally arising norms have $$\tau_{C^*} = 1$$, other naturally arising norms have $$\tau_{C^*} \geq 1/\sqrt{\vartheta}$$

- setting $$\delta = \varepsilon$$, we obtain $$\varepsilon \leq R_{\varepsilon}^P \cdot r_D^\varepsilon \leq 2\varepsilon$$, showing these two measures are inversely proportional (to within a factor of 2)

- exchanging the roles of $$P$$ and $$D$$
Application: Robust Optimization [J. Vera]

Amended format:

\[
P : \quad z^*(b) := \max_x \quad c^T x \quad \quad \quad \quad \quad \quad \quad D : \quad \min_y \quad b^T y
\]

s.t. \quad b - Ax \in K

s.t. \quad A^T y = c

\quad y \in K^*

For a given tolerance \( \varepsilon > 0 \), what is the limit on the size of a perturbation \( \Delta b \) so that \( |z^*(b + \Delta b) - z^*(b)| \leq \varepsilon \)?
Application: Robust Optimization, continued

\[ P : \quad z^*(b) := \max_x c^T x \quad \text{subject to} \quad b - Ax \in K \]
\[ D : \quad \min_y b^T y \quad \text{subject to} \quad A^T y = c \quad y \in K^* \]

**Theorem [Vera]:** Let \( \varepsilon > 0 \) and \( \Delta b \) satisfy:

\[ \|\Delta b\| \leq \tau_K \left( \frac{\varepsilon}{R^D_{\varepsilon}} \right) . \]

Then \( |z^*(b + \Delta b) - z^*(b)| \leq \varepsilon \).

The result says that \( \tau_K \cdot \varepsilon/R^D_{\varepsilon} \) is the required bound on the perturbation of the RHS needed to achieve a change of no more than \( \varepsilon \) in the value of the problem.
Some Relations with Renegar’s Condition
Number

For $\varepsilon \leq \|c\|_*$ it holds that:

$$R_{\varepsilon}^P \leq C^2(d) + C(d) \frac{\varepsilon}{\|c\|_*}$$

$$r_{\varepsilon}^P \geq \frac{\varepsilon \tau C}{3\|c\|_* (C^2(d) + C'(d))}$$
Natural Norms for $C$ for which $\tau_{C^*} = 1$

Given the regular cone $C$ and $w^0 \in \text{int} C^*$, the linear functional:

$$f(v) := (w^0)^T v$$

behaves like a norm when restricted to $v \in C$:

$f(v)$ is convex and positively homogeneous on $C$, and $f(v) > 0$ for $v \in C \setminus \{0\}$
Natural Norm Construction, continued

The natural norm that agrees with \( f(v) := (w^0)^T v \) on \( C \) is:

\[
\|v\|_{w^0} := \min_{v^1, v^2} (w^0)^T (v^1 + v^2)
\]

s.t. \( v^1 - v^2 = v \)

\( v^1 \in C \)
\( v^2 \in C \).

\[
\|v\|_{w^0} = (w^0)^T v \text{ for } v \in C
\]

We call this norm the “\( w^0 \)-norm” associated with the cone \( C \).
Properties of the $w^0$-norm

- If $v \in C$, then $\|v\|^{w^0} = (w^0)^T v$
- $\tau_{C^*} = 1$
- $\text{dist}_{w^0}^*(w, \partial C^*) \geq r \iff w - rw^0 \in C^*$
The $w^0$-norm Generalizes the $L_1$ norm

$$||v||^{w^0} := \min_{v^1, v^2} \ (w^0)^T (v^1 + v^2)$$

s.t. $v^1 - v^2 = v$
$v^1 \in C$
$v^2 \in C$.

$||v||^{w^0}$ is an exact generalization of the $L_1$ norm in the case when $C = IR^n_+$ and $w^0 = e$:

$$||v||_1 := \min_{v^1, v^2} \ e^T (v^1 + v^2)$$

s.t. $v^1 - v^2 = v$
$v^1 \in IR^n$
$v^2 \in IR^n_+$. 

37
The $w^0$-norm has Closed Form for Self-Scaled Cones

**Nonnegative Orthant:** $C = C^* = IR^n_+$, let $w^0 > 0$ be given.

$$\|v\|_{w^0} = \|W^0v\|_1 \text{ where } W^0 := \text{Diag}(w^0)$$

**Semidefinite Cone:** $C = C^* = S^{n \times n}_+$, let $W^0 \succ 0$ be given.

$$\|V\|_{W^0} = \left\|\lambda \left((W^0)^{1/2}V(W^0)^{1/2}\right)\right\|_1$$

**Second-Order Cone:** Closed form, but humorously complicated expression.
The $w^0$-norm for the Second-Order Cone

$C = C^* = Q^n := \{ v \in \mathbb{R}^n : \|(v_2, \ldots, v_n)\|_2 \leq v_1 \}$.

Rewrite $w^0 = (w^0_1, \bar{w})$ where $\bar{w} = (w^0_2, \ldots, w^0_n)$

Define: $M = \begin{pmatrix} w^0_1 & (\bar{w})^T \\ \bar{w} & (\sqrt{(w^0_1)^2 - \bar{w}^T \bar{w}}) I + \frac{\bar{w}\bar{w}^T}{w^0_1 + \sqrt{(w^0_1)^2 - \bar{w}^T \bar{w}}} \end{pmatrix}$

$\|v\|^{w^0} = \max\{\|(Mv)_1\|, \|(Mv)_2, \ldots, (Mv)_n\|_2 \}$.