

On Facial Reduction in the MOSEK Conic Optimizer

ICCOPT 2016, Tokyo, Japan

Henrik A. Friberg

www.mosek.com

The self-dual embedding





for a non-empty, closed, convex cone \mathcal{K} .



The self-dual embedding



 $\begin{array}{cccc} \mathsf{P}(\mathcal{K}): & \mathsf{D}(\mathcal{K}): \\ \text{minimize} & c^T x & \text{maximize} & b^T y \\ \text{subject to} & Ax = b, \\ & x \in \mathcal{K}, & \text{subject to} & c - A^T y = s, \\ & (s, y) \in \mathcal{K}^* \times \mathbb{R}^m, \end{array}$ $\mathsf{H}(\mathcal{K}):$

$$\begin{aligned} Ax - b\tau &= 0, \\ -A^T y - s + c\tau &= 0, \\ b^T y - c^T x - \kappa &= 0, \\ (x, s, y, \tau, \kappa) \in \mathcal{K} \times \mathcal{K}^* \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}_+. \end{aligned}$$

Observations

- Complementarity in $\mathbf{H}(\mathcal{K})$ by definition: $x^T s = \tau \kappa = 0$.
- $\tau > 0 \Rightarrow$ Complementary solution for $P(\mathcal{K})$ and $D(\mathcal{K})$.
- $\kappa > 0 \Rightarrow$ Improving ray for $\mathbf{P}(\mathcal{K})$ or $\mathbf{D}(\mathcal{K})$.

The self-dual embedding Third case: $\tau = \kappa = 0$



In Permenter, Friberg, and Andersen, *Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach*, 2015:

Ax = 0, $A^{T}y + s = 0,$ $b^{T}y - c^{T}x = 0,$ $(x, s, y, \tau, \kappa) \in \mathcal{K} \times \mathcal{K}^{*} \times \mathbb{R}^{m} \times 0 \times 0.$

The self-dual embedding Third case: $\tau = \kappa = 0$



In Permenter, Friberg, and Andersen, *Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach*, 2015:

$$\begin{array}{rcl} Ax &=& 0,\\ A^Ty+s &=& 0,\\ b^Ty-c^Tx &=& 0,\\ (x,s,y,\tau,\kappa)\in \mathcal{K}\times \mathcal{K}^*\times \mathbb{R}^m\times 0\times 0. \end{array}$$

Points like $(x, s, y, \tau, \kappa) = (0, 0, 0, 0, 0)$ will never be meaningful!

The self-dual embedding Third case: $\tau = \kappa = 0$



In Permenter, Friberg, and Andersen, *Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach*, 2015:

$$\begin{array}{rcl} Ax &=& 0,\\ A^Ty+s &=& 0,\\ b^Ty-c^Tx &=& 0,\\ (x,s,y,\tau,\kappa)\in \mathcal{K}\times \mathcal{K}^*\times \mathbb{R}^m\times 0\times 0. \end{array}$$

Key note from Theorem 1

Considering relint $\mathbf{H}(\mathcal{K})$, then $\tau = \kappa = 0$ is exceptional.

- If $\mathbf{P}(\mathcal{K})$ is troublesome: $\mathcal{K} \to \mathcal{K} \cap s^{\perp}$ (facial reduction).
- If $\mathbf{D}(\mathcal{K})$ is troublesome: $\mathcal{K}^* \to \mathcal{K}^* \cap x^{\perp}$ (facial reduction).



The self-dual embedding method can solve all instances!

Assuming oracle for relint $H(\mathcal{K})$ points.

In theory:

Yes, if tracking the central path with infinite precision.

In theory:

Yes, if tracking the central path with infinite precision.

In practice:

Issues in exceptional $\tau = \kappa = 0$ case:

• Lack of strict complementarity

 \Rightarrow slower and less accurate convergence.

- Ill-posedness (i.e., instability to infinitesimal perturbations)
 ⇒ overlapping termination criteria.
- Interior-point method
 - \Rightarrow when are τ and κ small enough?
 - \Rightarrow have to be careful about orthogonal complements,

e.g., when computing $\mathcal{K} \cap s^{\perp}$.





$$c^T x = 0, \quad A x = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$



$$c^T x = 0, \quad A x = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$



$$c^{\mathsf{T}}x = 0, \quad Ax = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$



$$c^T x = 0, \quad A x = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$



In theory:

$$c^T x = 0, \quad Ax = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$

This implies $x^{T}(c - A^{T}y) = x^{T}s \iff x^{T}s = 0$ and thus

$$s \in \mathcal{K}^* \longrightarrow s \in \mathcal{K}^* \cap x^\perp$$



In theory:

$$c^T x = 0, \quad A x = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$

This implies $x^{T}(c - A^{T}y) = x^{T}s \iff x^{T}s = 0$ and thus

 $s \in \mathcal{K}^* \quad \longrightarrow \quad s \in \mathcal{K}^* \cap x^{\perp}$ (a proper face of \mathcal{K}^*)



In theory:

$$c^T x = 0, \quad A x = 0, \quad x \in \mathcal{K} \setminus (\mathcal{K}^*)^{\perp}.$$

This implies $x^{T}(c - A^{T}y) = x^{T}s \iff x^{T}s = 0$ and thus

 $s \in \mathcal{K}^* \ \longrightarrow \ s \in \mathcal{K}^* \cap x^{\perp}$ (a proper face of \mathcal{K}^*)

In practice:

- Have $x \in \mathcal{K}$.
- Need negligible $|c^T x|$ and ||Ax||.
- Need significant $dist(x, (\mathcal{K}^*)^{\perp})$. (= ||x|| for proper cones)



In theory:

$$b^{\mathsf{T}}y = 0, \quad A^{\mathsf{T}}y + s = 0, \quad s \in \mathcal{K}^* \setminus \mathcal{K}^{\perp}$$

This implies $y^T A x = y^T b \iff s^T x = 0$ and thus

 $x \in \mathcal{K} \longrightarrow x \in \mathcal{K} \cap s^{\perp}$ (a proper face of \mathcal{K})

In practice:

- Have $s \in \mathcal{K}^*$.
- Need negligible $|b^T y|$ and $||A^T y + s||$.
- Need significant $dist(s, (\mathcal{K}^*)^{\perp})$. (= ||s|| for proper cones)



In theory:

$$b^{\mathsf{T}}y = 0, \quad A^{\mathsf{T}}y + s = 0, \quad s \in \mathcal{K}^* \setminus \mathcal{K}^{\perp}$$

This implies $y^T A x = y^T b \iff s^T x = 0$ and thus

 $x \in \mathcal{K} \longrightarrow x \in \mathcal{K} \cap s^{\perp}$ (a proper face of \mathcal{K})

In practice:

- Have $s \in \mathcal{K}^*$.
- Need negligible $|b^T y|$ and $||A^T y + s||$.
- Need significant $dist(s, (\mathcal{K}^*)^{\perp})$.

 $(\approx \|y\|$ significant)











If not satisfied, we may get $||A^Ty + s|| = 0$ for significant ||y|| and negligible ||s||.



If not satisfied, we may get $||A^Ty + s|| = 0$ for significant ||y|| and negligible ||s||. But then:

Linear row dependency certificate of $P(\mathcal{K})$

$$y^T A = 0, \quad y^T b = 0, \quad y \neq 0$$

Our proposal

In principle:

- **1** Negligible: ||Ax|| and $|c^Tx|$.
- **2** Negligible: $||A^Ty + s||$ and $|b^Ty|$.
- **3** Significant: ||x|| or ||y||.

In implementation:

1 max
$$(||Ax||, |c^Tx|) \le \epsilon \max(1, ||x||).$$

2 max $(||A^Ty + s||, |b^Ty|) \le \epsilon \max(1, ||y||)$
3 max $(||Ax||, |c^Tx|) < \epsilon ||x||$ or
max $(||A^Ty + s||, |b^Ty|) < \epsilon ||y||.$





H. Waki, M. Nakata and M. Muramatsu, "Strange Behaviors of Interior-point Methods for Solving Semidefinite Programming Problems in Polynomial Optimization", 2012:

<u>unboundDim1R*</u>

- Dual is SDP-relaxation of POP (minimized).
- Primal ill-posed with (p;d) = (0.0; 0.0^+).





H. Waki, M. Nakata and M. Muramatsu, "Strange Behaviors of Interior-point Methods for Solving Semidefinite Programming Problems in Polynomial Optimization", 2012:

<u>unboundDim1R*</u>

- Dual is SDP-relaxation of POP (minimized).
- Primal ill-posed with (p;d) = (0.0; 0.0^+).
- Higher-order Lasserre hierarchy means smaller perturbation exist after which (p;d) = (1.0; 1.0).



H. Waki, M. Nakata and M. Muramatsu, "Strange Behaviors of Interior-point Methods for Solving Semidefinite Programming Problems in Polynomial Optimization", 2012:

<u>unboundDim1R*</u>

- Dual is SDP-relaxation of POP (minimized).
- Primal ill-posed with (p;d) = (0.0; 0.0^+).
- Higher-order Lasserre hierarchy means smaller perturbation exist after which (p;d) = (1.0; 1.0).

*	Status	
1	Complementary solution of value 0.0	
2	Near to complementary solution of value 0.0 $>$	
3-4	Ill-posed certificate	
5-10	Complementary solution of value 1.0	



 ITE PFEAS
 DFEAS
 GFEAS
 MU

 0
 1.0e+00
 1.0e+00
 1.0e+00

 1
 3.1e-02
 3.1e-02
 1.1e+00

 2
 4.8e-05
 4.8e-05
 2.4e-02
 4.8e-05

 3
 1.1e-09
 1.1e-09
 2.2e-09
 1.1e-09

 Interior-point optimizer terminated. Time: 0.00.
 0.00.

Problem status : PRIMAL_AND_DUAL_FEASIBLE Solution status : OPTIMAL Primal. obj: -1.3336443272e-09 nrm: 1e+00 con: 1e-09 var: 0e+00 barvar: 2e-20 Dual. obj: 7.0491750159e-10 nrm: 2e+00 con: 0e+00 var: 1e-09

barvar: 1e-09



ITE PFEAS DFEAS GFEAS MU 2.0e+00 2.0e+00 1.0e+00 1.0e+00 0 1 2.9e-01 2.9e-01 1.7e-01 1.5e-01 2 5.2e-02 5.2e-02 7.2e-02 2.6e-02 . . . 37 4.1e-21 5.4e-20 4.7e-19 2.0e-21 38 4.1e-21 5.4e-20 4.7e-19 2.0e-21 4.1e-21 5.4e-20 4.7e-19 2.0e-21 39 Interior-point optimizer terminated. Time: 0.01.

Problem	status	:	PRIMAL_A	ND_DUAL	_FEASI	IBLE				
Solution	status	5:	NEAR_OPT	IMAL						
Primal.	obj: 2	2.12	14425554	e-04	nrm:	1e+00	con:	3e-13	barvar:	0e+00
Dual.	obj: 1	1.92	85790417	e-04	nrm:	3e+08	con:	0e+00	barvar:	3e-01



ITE PFEAS DFEAS GFEAS MU 0 2.0e+00 2.0e+00 1.0e+00 1.0e+00 1 2.9e-01 2.9e-01 1.7e-01 1.5e-01 2 5.2e-02 5.2e-02 7.2e-02 2.6e-02 . . . 37 4.1e-21 5.4e-20 4.7e-19 2.0e-21 38 4.1e-21 5.4e-20 4.7e-19 2.0e-21 4.1e-21 5.4e-20 4.7e-19 2.0e-21 39 Interior-point optimizer terminated. Time: 0.01.

Problem	status	:	PRIMAL_AND_DUAL	L_FEAS	IBLE				
Solution	status	:	NEAR_OPTIMAL						
Primal.	obj: 2	.12	214425554e-04	nrm:	1e+00	con:	3e-13	barvar:	0e+00
Dual.	obj: 1	. 92	285790417e-04	nrm:	3e+08	con:	0e+00	barvar:	3e-01



 ITE PFEAS
 DFEAS
 GFEAS
 MU

 0
 2.0e+00
 2.0e+00
 1.0e+00
 1.0e+00

 1
 3.3e-01
 3.3e-01
 1.5e-01
 1.7e-01

 2
 5.5e-02
 5.5e-02
 6.7e-02
 2.8e-02

 ...
 30
 7.9e-17
 2.7e-15
 6.6e-16
 4.0e-17

 31
 7.9e-17
 2.7e-15
 6.6e-16
 4.0e-17

 32
 7.9e-17
 2.7e-15
 6.6e-16
 4.0e-17

 Interior-point optimizer terminated.
 Time: 0.01.
 1.0e-10

Interior-point solution summary
Problem status : ILL_POSED
Solution status : PRIMAL_ILLPOSED_CER
Dual. obj: 1.3716138286e-08 nrm: 5e+00 con: 0e+00 barvar: 1e-07



H. Waki: "How to generate weakly infeasible semidefinite programs via Lasserre's relaxations for polynomial optimization", 2012:

CompactDim2R*

- Dual is SDP-relaxation of POP (minimized).
- Primal ill-posed with (p;d) = $(-\infty; -\infty^+)$.



H. Waki: "How to generate weakly infeasible semidefinite programs via Lasserre's relaxations for polynomial optimization", 2012:

CompactDim2R*

- Dual is SDP-relaxation of POP (minimized).
- Primal ill-posed with (p;d) = $(-\infty; -\infty^+)$.
- Higher-order Lasserre hierarchy means smaller perturbation exist after which (p;d) = (-1.5; -1.5).

*	Status
1	Primal infeasibility certificate
2	III-posed certificate
3-10	Complementary solution of value -1.5



Problem status : PRIMAL_INFEASIBLE Solution status : PRIMAL_INFEASIBLE_CER Dual. obj: -4.2224567689e-05 nrm: 3e+00 con: 0e+00 var: 2e-10 barvar: 3e-10



TTE PFEAS DFEAS GFEAS MU 0 2.0e+00 1.0e+00 2.0e+00 1.0e+00 1 4.2e-01 2.1e-01 4.2e-01 2.1e-01 2 7.1e-02 3.6e-02 5.3e-02 3.6e-02 . . . 23 1.2e-13 8.0e-14 4.9e-15 6.2e-14 24 2.6e-14 1.6e-14 3.1e-16 1.3e-14 25 3.5e-15 3.1e-15 5.3e-17 1.8e-15 Interior-point optimizer terminated. Time: 0.00.

Problem status : PRIMAL_INFEASIBLE Solution status : PRIMAL_INFEASIBLE_CER Dual. obj: -4.2224567689e-05 nrm: 3e+00 con: 0e+00 var: 2e-10 barvar: 3e-10



 ITE PFEAS
 DFEAS
 GFEAS
 MU

 0
 2.0e+00
 1.5e+00
 1.0e+00

 1
 4.4e-01
 3.3e-01
 3.8e-01
 2.2e-01

 2
 9.2e-02
 6.9e-02
 2.3e-01
 4.6e-02

 ...
 2
 1.8e-15
 3.6e-15
 5.0e-16
 9.0e-16

 26
 1.8e-15
 3.6e-15
 5.0e-16
 9.0e-16

 27
 1.8e-15
 3.6e-15
 5.0e-16
 9.0e-16

 Interior-point optimizer terminated. Time: 0.01.
 0.01.
 0.01.

Problem status : ILL_POSED Solution status : PRIMAL_ILLPOSED_CER Dual. obj: -1.7098776698e-06 nrm: 7e+00 con: 0e+00 barvar: 1e-08











Liu and Pataki, "Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming", 2015:

n=10, m=10	Strongly	/weakly	Weakly infeasible		
	infeasibl	e mix			
	Clean	Messy	Clean	Messy	
SEDUMI	87	27	0	0	
SDPT3	10	5	0	0	
MOSEK 7-series	63	17	0	0	
MOSEK 8.0.0.33(BETA)	98 w89	65 w64	100 w100	0	
PP+SEDUMI	100	27	100	0	



Liu and Pataki, "Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming", 2015:

n=10, m=20	Strongly	/weakly	Weakly infeasible		
	infeasibl	e mix			
	Clean	Messy	Clean	Messy	
SEDUMI	100	100	1	0	
SDPT3	100	96	0	0	
MOSEK 7-series	100	100	11	0	
MOSEK 8.0.0.33(BETA)	100 w1	69 w8	98 w98	0	
PP+SEDUMI	0	100	0	0	