Elimination of Free Variables for Solving ~~Semidefinite Programs~~ Efficiently

LOPs over Symmetric Cones

SIAM Conference on Optimization
Boston, May 10 - 13, 2008

Masakazu Kojima, Tokyo Institute of Technology

Sunyoung Kim, Ewha W. University

Hayato Waki, The University of Electro-Communications

Outline

1. Linear Optimization Problems (LOPs) with free variables
2. Elimination of free variables
3. Numerical results
4. Concluding remarks

Outline

1. Linear Optimization Problems (LOPs) with free variables
2. Elimination of free variables
3. Numerical results
4. Concluding remarks

Primal LOP with free vector variable $\boldsymbol{z}$

$$\mathcal{P}: \quad \begin{array}{ll} \text{min} & \boldsymbol{d}^T\boldsymbol{z} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \ \boldsymbol{x} \in \mathcal{K}, \end{array} \quad \begin{array}{l} \boldsymbol{D} : m \times k, \ \text{rank } \boldsymbol{D} = k, \\ \boldsymbol{A} : m \times n, \ \text{where } m \geq k \end{array}$$

Dual LOP with equality constraints

$$\mathcal{D}: \quad \text{max} \quad \boldsymbol{b}^T\boldsymbol{y} \quad \text{s.t.} \quad \boldsymbol{D}^T\boldsymbol{y} = \boldsymbol{d}, \ \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y} \in \mathcal{K}.$$

Here $\mathcal{K}$: a symmetric cone such as SDP, SOCP, LP cones and their products.

- How to handle free variables is an important issue in primal-dual interior-point methods for SDPs.
- Some methods have been developed:
  - (a) free $z = z_+ - z_-$, $z_+ \geq 0, \ z_- \geq 0$
  - (b) a second order cone
  - (c) a regularization technique of Meszaros by Anjos-Burer
  - (d) elimination of the free var. $\boldsymbol{z}$ and the eq. $\boldsymbol{D}^T\boldsymbol{y} = \boldsymbol{d}$ by Kobayashi-Nakata-Kojima — this talk

Primal LOP with free vector variable $\boldsymbol{z}$

$$\mathcal{P}: \quad \begin{array}{ll} \min & \boldsymbol{d}^T\boldsymbol{z} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \ \boldsymbol{x} \in \mathcal{K}, \end{array} \quad \begin{array}{l} \boldsymbol{D} : m \times k, \ \text{rank } \boldsymbol{D} = k, \\ \boldsymbol{A} : m \times n, \ \text{where } m \geq k \end{array}$$

Dual LOP with equality constraints

$$\mathcal{D}: \quad \max \quad \boldsymbol{b}^T\boldsymbol{y} \quad \text{s.t.} \quad \boldsymbol{D}^T\boldsymbol{y} = \boldsymbol{d}, \ \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y} \in \mathcal{K}.$$

Here $\mathcal{K}$: a symmetric cone such as SDP, SOCP, LP cones and their products. Two (equivalent) approaches:

- Primal approach: Eliminate free variable $\boldsymbol{z}$ by pivoting or LU factorization $\Rightarrow$

$$\widehat{\mathcal{P}}: \quad \begin{array}{ll} \min & \hat{\boldsymbol{c}}^T\boldsymbol{x} + \hat{\gamma} \\ \text{s.t.} & \widehat{\boldsymbol{A}}_2\boldsymbol{x} = \hat{\boldsymbol{b}}_2, \ \boldsymbol{x} \in \mathcal{K}, \end{array} \quad \widehat{\boldsymbol{A}}_2 : (m-k) \times n.$$

- Dual : Solve $\boldsymbol{D}^T\boldsymbol{y} = \boldsymbol{d}$ in $\boldsymbol{y}_1$, $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2) \in \mathbb{R}^{k+(m-k)} \quad \Rightarrow$

$$\widehat{\mathcal{D}}: \quad \max \quad \hat{\boldsymbol{b}}_2^T\boldsymbol{y}_2 + \hat{\gamma} \quad \text{s.t.} \quad \hat{\boldsymbol{c}} - \widehat{\boldsymbol{A}}_2^T\boldsymbol{y}_2 \in \mathcal{K}.$$

- The size gets smaller, but $\widehat{\boldsymbol{A}}_2$ could get denser than $\boldsymbol{A}$.
- Numerical stability in pivoting or solving $\boldsymbol{D}^T\boldsymbol{y} = \boldsymbol{d}$ in $\boldsymbol{y}_1$.

Primal LOP with free vector variable $\boldsymbol{z}$

$$\mathcal{P}: \quad \begin{array}{ll} \text{min} & \boldsymbol{d}^T\boldsymbol{z} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathcal{K}, \end{array} \qquad \begin{array}{l} \boldsymbol{D} : m \times k, \text{ rank } \boldsymbol{D} = k, \\ \boldsymbol{A} : m \times n, \text{ where } m \geq k \end{array}$$

A stable sparse LU factorization to $\boldsymbol{D}$ (Markowitz pivot selection and its variations) — for simplicity,

$$\boldsymbol{P}\boldsymbol{D}\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{U} \quad \text{or} \quad \boldsymbol{D} = \boldsymbol{P}^T\boldsymbol{L}\boldsymbol{U}\boldsymbol{Q}^T \qquad = \boldsymbol{L}\boldsymbol{U}$$

$$k, \qquad \boldsymbol{U} \ : \ k \times k \text{ upper triangular,}$$

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{L}_1 \\ \boldsymbol{L}_2 \end{pmatrix} \begin{array}{c} k \\ m-k \end{array}, \quad \boldsymbol{L}_1 \ : \ \text{lower triangular,}$$

$\boldsymbol{P}$: an $m \times m$ permutation matrix, $= \boldsymbol{I}$

$\boldsymbol{Q}$: a $k \times k$ permutation matrix, $= \boldsymbol{I}$

Primal LOP with free vector variable $\boldsymbol{z}$

$$\mathcal{P}: \quad \begin{array}{ll} \text{min} & \boldsymbol{d}^T\boldsymbol{z} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathcal{K}, \end{array} \qquad \begin{array}{l} \boldsymbol{D} : m \times k,\ \text{rank } \boldsymbol{D} = k, \\ \boldsymbol{A} : m \times n,\ \text{where } m \geq k \end{array}$$

A stable sparse LU factorization to $\boldsymbol{D}$

(Markowitz pivot selection and its variations)

$$\boldsymbol{D} = \boldsymbol{L}\boldsymbol{U}$$

$$k, \qquad \boldsymbol{U} : k \times k \text{ upper triangular},$$

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{L}_1 \\ \boldsymbol{L}_2 \end{pmatrix} \begin{array}{c} k \\ m-k \end{array}, \quad \boldsymbol{L}_1 : \text{ lower triangular},$$

$$\widehat{\mathcal{P}}: \quad \begin{array}{ll} \text{min} & \hat{\boldsymbol{c}}^T\boldsymbol{x} + \hat{\gamma} \\ \text{s.t.} & \widehat{\boldsymbol{A}}_2\boldsymbol{x} = \hat{\boldsymbol{b}}_2,\ \boldsymbol{x} \in \mathcal{K},\ \boldsymbol{z} = \boldsymbol{U}^{-1}(\hat{\boldsymbol{b}}_1 - \widehat{\boldsymbol{A}}_1\boldsymbol{x}). \end{array} \qquad \widehat{\boldsymbol{A}}_2 : (m-k) \times n,\ \widehat{\boldsymbol{A}}_1 : k \times n$$

$$\hat{\boldsymbol{c}} = \boldsymbol{c} - \widehat{\boldsymbol{A}}_1^T \boldsymbol{U}^{-T}\boldsymbol{d},\ \hat{\gamma} = \hat{\boldsymbol{b}}_1^T \boldsymbol{U}^{-T}\boldsymbol{d},$$

$$\begin{pmatrix} \widehat{\boldsymbol{A}}_1 \\ \widehat{\boldsymbol{A}}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_1 & \boldsymbol{O} \\ \boldsymbol{L}_2 & \boldsymbol{I} \end{pmatrix}^{-1} \boldsymbol{A},\ \begin{pmatrix} \hat{\boldsymbol{b}}_1 \\ \hat{\boldsymbol{b}}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_1 & \boldsymbol{O} \\ \boldsymbol{L}_2 & \boldsymbol{I} \end{pmatrix}^{-1} \boldsymbol{b},$$

Primal LOP with free vector variable $\boldsymbol{z}$

$$\mathcal{P}: \quad \begin{array}{ll} \min & \boldsymbol{d}^T\boldsymbol{z} + \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathcal{K}, \end{array} \qquad \begin{array}{l} \boldsymbol{D} : m \times k,\ \text{rank}\ \boldsymbol{D} = k, \\ \boldsymbol{A} : m \times n,\ \text{where}\ m \geq k \end{array}$$

A stable sparse LU factorization to $\boldsymbol{D}$
(Markowitz pivot selection and its variations)

$$\boldsymbol{D} = \boldsymbol{L}\boldsymbol{U}$$

$$k, \qquad \boldsymbol{U} : k \times k \text{ upper triangular},$$

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{L}_1 \\ \boldsymbol{L}_2 \end{pmatrix} \begin{array}{c} k \\ m-k \end{array},\ \boldsymbol{L}_1 : \text{ lower triangular},$$

$$\widehat{\mathcal{P}}: \quad \begin{array}{ll} \min & \hat{\boldsymbol{c}}^T\boldsymbol{x} + \hat{\gamma} \\ \text{s.t.} & \widehat{\boldsymbol{A}}_2\boldsymbol{x} = \hat{\boldsymbol{b}}_2,\ \boldsymbol{x} \in \mathcal{K},\ \boldsymbol{z} = \boldsymbol{U}^{-1}(\hat{\boldsymbol{b}}_1 - \widehat{\boldsymbol{A}}_1\boldsymbol{x}). \end{array} \qquad \widehat{\boldsymbol{A}}_2 : (m-k) \times n,\ \widehat{\boldsymbol{A}}_1 : k \times n$$

- k is larger $\Rightarrow$ smaller size
- $\boldsymbol{LU}$ factorization is well-conditioned $\Rightarrow$ higher accuracy
- $\boldsymbol{LU}$ factorization (or $\widehat{\boldsymbol{A}}_2$) is sparser $\Rightarrow$ more efficient

A stable sparse LU factorization (Markowitz pivot selection):

$$\begin{pmatrix} * & * & * & * & * & * & \\ 0 & * & * & * & * & * & \\ 0 & 0 & * & * & * & * & \\ 0 & 0 & * & * & * & * & i \\ 0 & 0 & * & * & * & * & \\ 0 & 0 & * & * & * & * & \\ & & & & \mathbf{j} & & \end{pmatrix}$$

1. Let a_{ij} $((i,j) \in \Lambda_0)$ be all candidates of pivot elements.

2. Numerical stability: Let

$$\Lambda_1 = \{(i,j) \in \Lambda_0 : |a_{ij}| \geq \rho \max_{i';\,(i',\mathbf{j}) \in \Lambda_0} |a_{i'j}|\}.$$

Here $\rho \in (0,1)$ denotes a parameter; stablity $\uparrow$ as $\rho \to 1$.

3. Sparsity: Let

$$\begin{aligned}(i^*, j^*) \quad = \quad & \text{arg min}_{(i,j) \in \Lambda_1} \\ & (\text{\# of nonzeros in } i\text{th row} - 1) \times \\ & (\text{\# of nonzeros in } j\text{th column} - 1)\end{aligned}$$

4. Choose (i^*, j^*) as the pivot element.

- $\rho = 0.25$ in the next section "Numerical results".

Outline

1. Linear Optimization Problems (LOPs) with free variables
2. Elimination of free variables
3. Numerical results
4. Concluding remarks

Numerical results

- SDPA, conversion + SDPA, SeDuMi, conversion+SeDuMi, sqlp(SDPT3), HSDsqlp(SDPT3).
- 49 SDP relaxation problems of POPs (from global lib), sensor network localization problems, ODEs and PDEs.
- SDPA was able to solve 42 SDPs among 49 with high accuracy, $\log_{10}$(duality gap) ≤ -5.
- 8 SDPs showed worse duality gaps in SDPA and/or conversion + SDPA among 49.
- conversion + SDPA was not effective for 2 SDPs among 8 SDPs .

$\Longrightarrow$ Numerical results on only 8 SDPs

Numerical results on 8 SDPs

	$\log_{10}$(duality gap)					
8 SDPs	SDPA	conv+ SDPA	sedumi	conv+ sedumi	SDPT3 sqlp	SDPT3 HSD
ex9_2_3	+0.01	-7.28	-10.14	-8.00	-5.16	-5.79
st_e42	+1.03	-5.11	-7.24	-6.72	-4.18	-8.71
mhw4d	-1.41	-5.87	-9.06	-8.40	-9.19	-9.12
EQD20	+3.23	-12.53	-5.47	-9.54	-4.40	-2.08
odeT100	-0.43	-7.19	-2.71	-7.86	-5.91	-12.03
odeT500	+1.56	-6.97	-3.02	-7.07	-1.96	-2.09
ex5_3_2	-8.30	-2.21	-4.77	-6.72	-0.12	-5.99
alkylation	+5.97	+5.37	-0.00	-0.00	+0.00	-0.07

Low accuracy High accuracy

Numerical results on 8 SDPs

$$\text{p.f} = \|\boldsymbol{D}\boldsymbol{z} + \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_{\text{inf}}$$

	$\log_{10}$(p.f)					
		conv+		conv+	SDPT3	
8 SDPs	SDPA	SDPA	sedumi	sedumi	sqlp	HSD
ex9_2_3	-2.84	-12.48	-13.59	-11.13	-5.22	-3.56
st_e42	-2.48	-5.67	-8.41	-7.90	-5.27	-5.09
mhw4d	-5.37	-7.43	-9.57	-9.03	-10.08	-8.92
EQD20	+2.85	-7.43	-6.00	-6.93	-7.18	-4.57
odeT100	-4.64	-7.27	-8.89	-8.22	-5.02	-7.01
odeT500	-5.90	-5.30	-8.91	-7.52	-1.64	-0.43
ex5_3_2	-7.64	-6.77	-8.62	-8.93	+0.19	-4.43
alkylation	+0.29	-0.02	-0.00	-1.26	-0.00	-0.00

Low accuracy High accuracy

Numerical results on 8 SDPs

d.f $=$max $\{|$min. { eigenvalues of $\boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y},\ 0\}|,\ \|\boldsymbol{D}^T\boldsymbol{y} - \boldsymbol{d}\|_{\text{inf}}\}$

	$\log_{10}$(d.f)					
		conv+		conv+	SDPT3	
8 SDPs	SDPA	SDPA	sedumi	sedumi	sqlp	HSD
ex9_2_3	-7.11	-7.04	-11.72	-9.90	-5.87	-8.88
st_e42	-6.99	-7.18	-9.60	-15.08	-9.53	-13.47
mhw4d	-7.15	-7.61	-10.04	-8.52	-12.21	-11.59
EQD20	-0.84	-7.16	-6.73	-9.75	-6.36	-5.07
odeT100	-7.22	-17.22	-8.88	-8.97	-17.70	-17.28
odeT500	-4.91	-17.70	-7.99	-7.38	-10.61	-8.99
ex5_3_2	-7.10	-7.23	-8.36	-8.59	-5.87	-8.88
alkylation	-0.77	-0.43	-0.04	+0.73	+0.11	-0.69

Low accuracy High accuracy

Outline

1. Linear Optimization Problems (LOPs) with free variables
2. Elimination of free variables
3. Numerical results
4. Concluding remarks

1. Solving SDPs with free variables is not easy.
2. A numerical method for eliminating free variables for SDPs.
3. The basic idea behind the method is quite natural, and it works effectively in some extent but not enough sometime.
4. Further study.