Semidefinite Programming
Lecture 7

scribe: Yuming Zhang

March 20 , 1995

1 Potential Functions and Projective Transformations

In this lecture, we develope a potential reduction method for the primal problem so that , within $O(\sqrt{n}|\log \epsilon|)$ iterations, we get an approximate solution with at least ϵ relative accuracy, if ϵ is sufficiently small. Our development closely follows Ye's projective algorithm for linear programming . Ye's complexity analysis is also extended to semidefinite programs.

Let's recall some definitions in section 6.4

Let $q > 0$, and $\underline{z}$ be a given constant known to be a lower bound on the optimal value z^* of the primal problem. Let X be an interior primal feasible matrix, $\mathbf{y}$ an interior dual feasible vector, and $S{:=}C - \sum_{i=1}^{m} y_i A_i$; thus, $X \succ 0$ and $S \succ 0$.

The *Primal Potential Function*:

$$\phi(X, \underline{z}) = q \ln(C \bullet X - \underline{z}) - \ln \det X, \tag{1.1}$$

Primal-Dual Potential Function:

$$\psi(X, S) = q \ln(X \bullet S) - \ln \det(XS). \tag{1.2}$$

Before describing the algorithm we state the following lemma which is a direct generalization of a similar lemma that appears in the analysis of most interior point linear programming methods. (Recall that $\rho(X)$ is the spectral radius of matrix X, which equals its largest eigenvalue when X is positive semidefinite.)

LEMMA 1.1. *Let X be a symmetric $n \times n$ matrix. If $0 \prec X \prec I$, then*

$$\ln \det X \geq \text{trace } X - n - \frac{\text{trace } (X - I)^2}{2[1 - \rho(X - I)]}.$$

Proof. In most interior-point linear programming algorithms it is shown that if $\|\mathbf{x} - \mathbf{1}\|_\infty < 1$ then

$$\sum_{j=1}^{n} \ln x_j \geq (\mathbf{1}^T\mathbf{x} - n) - \frac{\|\mathbf{x} - \mathbf{1}\|^2}{2(1 - \|\mathbf{x} - \mathbf{1}\|_\infty)}$$

which is easily proved by expanding $\ln x$. Now to prove the lemma simply substitute $\lambda_j(X)$ for x_j. ■

Next, we are going to define *projective transformation* which is used to create a new iterate from the current iterate:

Let $X_0 \succ 0$ be our current interior primal feasible point. To find a symmetry preserving projective transformation that maps X_0 to the identity matrix I, let L_0 be any $n \times n$ matrix such that $L_0 L_0^T = X_0$. There are infinitely many choices for L_0. For instance, it could be a Cholesky factor of X_0, or it could be its square root, $X_0^{1/2}$. We shall see shortly that it does not matter how we select L_0 as it will not affect the algorithm's behavior and performance. For a fixed integer r, define $\mathcal{T} : S^{n\times n} \to S^{n\times n} \times \Re^r$ (here, S is the *semidefinite positive cone*), such that $\mathcal{T}(X) = (\overline{X}, \bar{\mathbf{x}})$. Then:

$$\overline{X} = \frac{(n+r)L_0^{-1} X L_0^{-T}}{r + X_0^{-1} \bullet X} \text{ and } \bar{\mathbf{x}} = \left(\frac{n+r}{r + X_0^{-1} \bullet X} \right) \mathbf{1}. \tag{1.3}$$

We call $\mathcal{T}$ *projective transformation.*

Its inverse transformation is defined as:

$$X = \mathcal{T}^{-1}(\overline{X}, \bar{\mathbf{x}}) = \frac{L_0 \overline{X} L_0^T}{\sum \bar{x}_j / r}. \tag{1.4}$$

Remark: This inverse transformation is different from the usual definition.

Under $\mathcal{T}$, the primal SDP problem is transformed into

$$\begin{array}{ll} \min & \overline{C} \bullet \overline{X} + \bar{\mathbf{c}}(\underline{z})^T \bar{\mathbf{x}} \\ \text{s.t.} & \bar{\mathcal{A}} \mathbf{vec}\, \overline{X} + \bar{A} \bar{\mathbf{x}} = \mathbf{0} \\ & \operatorname{trace} \overline{X} + \mathbf{1}^T \bar{\mathbf{x}} = n + r \\ & \overline{X} \succeq 0 \\ & \bar{\mathbf{x}} \geq \mathbf{0}, \end{array} \tag{1.5}$$

where

$$\overline{C} = L_0^T C L_0;\ \bar{\mathbf{c}}(\underline{z}) = -(\underline{z}/r)\,\mathbf{1};\ \bar{A}_i = L_0^T A_i L_0;\ \bar{\mathcal{A}} = \mathcal{A}(L_0 \otimes L_0);\ \bar{A} = (-1/r)\,\mathbf{b}\mathbf{1}^T.$$

Note that $\bar{A}$ is an $m \times r$ rank one matrix. The transformed problem may be viewed as a mixed linear and semidefinite program. We may define the following primal potential function for the transformed problem:

$$\bar{\phi}(\overline{X}, \bar{\mathbf{x}}, \underline{z}) = q \ln \left[\overline{C} \bullet \overline{X} + \bar{\mathbf{c}}(\underline{z})^T \bar{\mathbf{x}} \right] - \ln \det \overline{X} - \sum_{j=1}^{r} \ln \bar{x}_j. \tag{1.6}$$

The following invariant property holds for the potential functions under projective transformations:

LEMMA 1.2. *If* $\bar{x}_1 = \cdots = \bar{x}_r$, *and* $q = n + r$ *and* $X = \mathcal{T}^{-1}(\overline{X}, \bar{\mathbf{x}})$ *then*

$$\phi(X, \underline{z}) - \phi(X_0, \underline{z}) = \overline{\phi}(\overline{X}, \bar{\mathbf{x}}, \underline{z}) - \overline{\phi}(I, \mathbf{1}, \underline{z}). \tag{1.7}$$

Also, the following result is easily proved by expanding $\overline{\phi}$ and applying lemma 1.1; later we use it to prove the reduction in the primal-dual potential function.

COROLLARY 1.1. *For* $q = n + r$ *we have*

$$\overline{\phi}(\overline{X}, \bar{\mathbf{x}}, \underline{z}) - \overline{\phi}(I, \mathbf{1}, \underline{z}) \le (n+r) \ln \left(\frac{\overline{C} \bullet \overline{X} + \bar{\mathbf{c}}(\underline{z})^T \bar{\mathbf{x}}}{\text{trace } \overline{C} + \bar{\mathbf{c}}(\underline{z})^T \mathbf{1}} \right) + \frac{\|\overline{X} - I\|^2 + \|\bar{\mathbf{x}} - \mathbf{1}\|^2}{2(1 - \|\overline{X} - I\| + \|\bar{\mathbf{x}} - \mathbf{1}\|)}.$$

2 A potential reduction algorithm

Similar to Karmarkar algorithm in linear programming, we replace the inequality constraints $\overline{X} \succeq 0$ and $\bar{\mathbf{x}} \ge \mathbf{0}$ by an inscribed "ball" constraint, but here the ball is centered at $(I, \mathbf{1})$. Therefore, we have the "ball optimization" problem:

$$\begin{array}{ll} \min & \overline{C} \bullet \overline{X} + \bar{\mathbf{c}}(\underline{z})^T \bar{\mathbf{x}} \\ \text{s.t.} & \bar{\mathcal{A}} \mathbf{vec}\, \overline{X} + \bar{A} \bar{\mathbf{x}} = \mathbf{0} \\ & \text{trace } \overline{X} + \mathbf{1}^T \bar{\mathbf{x}} = n + r \\ & \|X - I\|^2 + \|\mathbf{x} - \mathbf{1}\|^2 \le \beta^2 < 1 \end{array} \tag{2.8}$$

where β is a fixed constant between 0 and 1 to be determined shortly. Once we solve this problem and map the result back to the original space, we get a point that serves as a candidate for the next iterate. The solution of this problem is obtained as follows:

$$\begin{pmatrix} \mathbf{vec}\overline{X}_1 \\ \bar{\mathbf{x}}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{vec} I \\ \mathbf{1} \end{pmatrix} - \beta \frac{P(\underline{z})}{\|P(\underline{z})\|}, \tag{2.9}$$

and the candidate for the new primal iterate is given by:

$$X(\underline{z}) = \mathcal{T}^{-1}(\overline{X}_1, \bar{\mathbf{x}}_1), \tag{2.10}$$

where

$$P(\underline{z}) = \mathcal{P}_{\mathcal{A}'} \begin{pmatrix} \mathbf{vec}\overline{C} \\ \bar{\mathbf{c}}(\underline{z}) \end{pmatrix}, \quad \mathcal{A}' = \begin{pmatrix} \bar{\mathcal{A}} & \bar{A} \\ (\mathbf{vec}\, I)^T & \mathbf{1}^T \end{pmatrix}, \tag{2.11}$$

and $\mathcal{P}_{\mathcal{A}'}(\mathbf{u})$ is the projection of the (n^2+r)-vector $\mathbf{u}$ to the null space of $\mathcal{A}'$. After expansion, the projection $\mathcal{P}_{\mathcal{A}'}$ in (2.11) becomes:

$$P(\underline{z}) = \left(I - \frac{[\mathbf{vec}^T I | \mathbf{1}^T][\mathbf{vec}^T I | \mathbf{1}^T]^T}{n+r} \right) \left(I - [\bar{\mathcal{A}}|\bar{A}]^T ([\bar{\mathcal{A}}|\bar{A}][\bar{\mathcal{A}}|\bar{A}]^T)^{-1} [\bar{\mathcal{A}}|\bar{A}] \right) \begin{pmatrix} \mathbf{vec}\bar{C} \\ \bar{\mathbf{c}}(\underline{z}) \end{pmatrix}. \tag{2.12}$$

Define:

$$\begin{aligned} \mathbf{y}(\underline{z}) &= \left([\bar{\mathcal{A}}|\bar{A}][\bar{\mathcal{A}}|\bar{A}]^T \right)^{-1} [\bar{\mathcal{A}}|\bar{A}] \begin{pmatrix} \mathbf{vec}\overline{C} \\ \bar{\mathbf{c}}(\underline{z}) \end{pmatrix} \\ &= \left(\mathcal{A}(X_0 \otimes X_0)\mathcal{A}^T + (1/r)\, \mathbf{b}\mathbf{b}^T \right)^{-1} [\mathcal{A}\mathbf{vec}(X_0 C X_0) + (\underline{z}/r)\, \mathbf{b}] \end{aligned} \tag{2.13}$$

and

$$S(\underline{z}) = C - \mathrm{Mat}(\mathcal{A}^T \mathbf{y}(\underline{z})). \tag{2.14}$$

$S(\underline{z})$ and $\mathbf{y}(\underline{z})$ serve as candidates for the new dual iterates. In terms of these quantities $P(\underline{z})$ may be written as:

$$P(\underline{z}) = \begin{pmatrix} \mathbf{vec}(L_0^T S(\underline{z}) L_0) \\ \frac{\mathbf{b}^T \mathbf{y}(\underline{z}) - \underline{z}}{r} \mathbf{1} \end{pmatrix} - \frac{C \bullet X_0 - \underline{z}}{n+r} \begin{pmatrix} \mathbf{vec} I \\ \mathbf{1} \end{pmatrix}. \tag{2.15}$$

Observe that $X(\underline{z})$, $S(\underline{z})$ and $\mathbf{y}(\underline{z})$ are all independent of L_0; in fact in actual computation we do not need to have L_0 explicitly.

The main result to be proved is that first, at least one of the following holds:

1. Either $X(\underline{z}) \succ 0$ and thus primal feasible, or
2. $S(\underline{z}) \succ 0$ and therefore $(\mathbf{y}(\underline{z}), S(\underline{z}))$ is dual feasible

Furthermore choosing either one of the feasible candidates reduces the value of the primal-dual potential function ψ by a constant amount. First observe that $\mathcal{P}_{\mathcal{A}'}$ is a projector, that is $\mathcal{P}_{\mathcal{A}'}^2 = \mathcal{P}_{\mathcal{A}'}$. Therefore, from (2.9) we get:

$$\overline{C} \bullet (\overline{X}_1 - I) + \bar{\mathbf{c}}(\underline{z})^T (\bar{\mathbf{x}} - \mathbf{1}) = -\beta \|P(\underline{z})\|$$

Hence, noting that $\ln(1+x) \le x$, for nonnegative x, corollary 1.1 implies:

COROLLARY 2.1. *Let $q = n + r$ and $\overline{X}_1$ and $\bar{\mathbf{x}}$ be as in (2.9). Then*

$$\overline{\phi}(\overline{X}, \bar{\mathbf{x}}, \underline{z}) - \overline{\phi}(I, \mathbf{1}, \underline{z}) \le -(n+r)\beta \frac{\|P(\underline{z})\|}{\mathbf{c}(\underline{z})^T \mathbf{1} + \operatorname{trace} \overline{C}} + \frac{\beta^2}{2(1-\beta)}.$$

Let Δ_0 be the size of the duality gap in the current iterate, that is

$$\Delta_0 = C \bullet X_0 - \underline{z}$$

and let

$$\Delta_1 = S(\underline{z}) \bullet X_0 = C \bullet X_0 - \mathbf{b}^T \mathbf{y}(\underline{z}).$$

Thus Δ_1 should be interpreted as the value of the duality gap if we prove the following lemma:

LEMMA 2.1. *If there is some real number α with $0 < \alpha < 1$, such that*

$$\|P(\underline{z})\| \le \alpha \frac{\Delta_0}{n+r}$$

then $S(\underline{z}) \succ 0$, and $\mathbf{b}^T \mathbf{y}(\underline{z}) > \underline{z}$. Furthermore,

$$\left\| L_0^T S(\underline{z}) L_0 - \frac{\Delta_1}{n} I \right\| \le \frac{\Delta_1}{n} \alpha \sqrt{\frac{n + n^2/r}{n + n^2/r - \alpha^2}}, \tag{2.16}$$

and

$$\left| \frac{n+r}{n} \frac{\Delta_1}{\Delta_0} - 1 \right| \le \frac{\alpha}{\sqrt{n + n^2/r}}. \tag{2.17}$$

Proof. Suppose $S(\underline{z}) \not\succ 0$. Then $L_0^T S(\underline{z}) L_0$ is not positive definite and so some of its eigenvalues are less than or equal to 0. Thus, from (2.15) we have

$$\|P(\underline{z})\| \geq \rho\left(\frac{\Delta_0}{n+r} I - L_0^T S(\underline{z}) L_0\right) \geq \frac{\Delta_0}{n+r},$$

a contradiction. Also, If $\mathbf{b}^T \mathbf{y}(\underline{z}) \leq \underline{z}$ then from (2.15) we have

$$\|P(\underline{z})\| \geq \frac{\Delta_0}{n+r} - \frac{\mathbf{b}^T \mathbf{y}(\underline{z}) - \underline{z}}{r} \geq \frac{\Delta_0}{n+r},$$

which is again a contradiction. Now from (2.15) we have

$$P(\underline{z}) = \left(\begin{array}{c} \left[\mathbf{vec}(L_0^T S(\underline{z}) L_0) - \frac{\Delta_1}{n} I\right] - \left[\frac{\Delta_0}{n+r} - \frac{\Delta_1}{n}\right] \mathbf{vec} I \\ \left[\frac{\Delta_0 - \Delta_1}{r} - \frac{\Delta_0}{n+r}\right] \mathbf{1} \end{array} \right).$$

Since $I \bullet [(L_0^T S(\underline{z}) L_0) - (\Delta_1/n) I] = 0$, we have

$$\begin{aligned} \|P(\underline{z})\|^2 &= \left\|L_0^T S(\underline{z}) L_0 - \frac{\Delta_1}{n} I\right\|^2 + n\left(\frac{\Delta_0}{n+r} - \frac{\Delta_1}{n}\right)^2 + r\left(\frac{\Delta_0 - \Delta_1}{r} - \frac{\Delta_0}{n+r}\right)^2 \\ &= \left\|L_0^T S(\underline{z}) L_0 - \frac{\Delta_1}{n} I\right\|^2 + \left(n + \frac{n^2}{r}\right)\left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r}\right)^2. \end{aligned} \tag{2.18}$$

If (2.16) is false then from (2.18) we have

$$\begin{aligned} \|P(\underline{z})\|^2 &> \left(\frac{\Delta_1}{n}\right)^2 \alpha^2 \frac{n + n^2/r}{n + n^2/r - \alpha^2} + \left(n + \frac{n^2}{r}\right)\left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r}\right)^2 \\ &\geq \alpha^2 \left(\frac{\Delta_0}{n+r}\right)^2. \end{aligned} \tag{2.19}$$

(The last inequality is proved by taking the right hand side of the first inequality as a quadratic function in Δ_1/n and minimizing it.) But (2.19) contradicts the assumption of the lemma, so (2.16) must be true. Finally, since (2.19) is false, we have:

$$\left(n + \frac{n^2}{r}\right)\left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r}\right)^2 \leq \alpha^2 \left(\frac{\Delta_0}{n+r}\right)^2,$$

from which (2.17) follows. ∎

Now we may prove the potential reduction theorem.

THEOREM 2.1. *Let X_0 be any interior feasible matrix for the primal problem and $\mathbf{y}_0$ interior feasible for the dual. Let also, $r = \lceil \sqrt{n}\, \rceil$ and $q = n + r$, $S_0 = C - \sum_{i=1}^m y_i A_i$, $\underline{z}_0 = \mathbf{b}^T \mathbf{y}_0$, $X(\underline{z}) = \mathcal{T}^{-1}(\overline{X}_1, \bar{\mathbf{x}}_1)$, as in (1.4), $\mathbf{y}_1 = \mathbf{y}(\underline{z}_0)$, and $S_1 = S(\underline{z}_0)$. Then there exist an absolute constant δ such that either*

$$\psi(X(\underline{z}), S_0) \leq \psi(X_0, S_0) - \delta,$$

or

$$\psi(X_0, S_1) \leq \psi(X_0, S_0) - \delta,$$

Furthermore, if we set $\alpha = 0.55$ and $\beta = 0.3$, then $\delta > 0.1$.

Proof. If for some constant $0 < \alpha < 1$

$$\|P(\underline{z})\| \geq \alpha \frac{\Delta_0}{n+r}$$

then

$$\begin{aligned}
\psi(X(\underline{z}), S_0) - \psi(X_0, S_0) &= \phi(X(\underline{z}), \underline{z}_0) - \phi(X_0, \underline{z}_0) \\
&= \overline{\phi}(\overline{X}(\underline{z}), \bar{\mathbf{x}}_1, \underline{z}_0) - \overline{\phi}(I, \mathbf{1}, \underline{z}_0) \\
&\leq -\beta\alpha + \frac{\beta^2}{2(1-\beta)}
\end{aligned}$$

(the last inequality is true by corollary 2.1). Otherwise, the conditions of lemma 2.1 are satisfied. Also applying lemma 1.1 to $(n/\Delta_1)L_0^T S_1 L_0$, and setting $\gamma = \alpha\sqrt{\frac{n+n^2/r}{n+n^2/r-\alpha^2}}$ we have:

$$\begin{aligned}
n \ln X_0 \bullet S_1 - \ln\det X_0 S_1 &= n \ln\left(\frac{nX_0 \bullet S_1}{\Delta_1}\right) - \ln\det \frac{nX_0S_1}{\Delta_1} \\
&= n \ln n - \ln\det \frac{nX_0S_1}{\Delta_1} \\
&\leq n \ln n + \frac{\|nL_0^T S_1 L_0/\Delta_1 - I\|^2}{2(1 - \|nL_0^T S_1 L_0/\Delta_1 - I\|)} \\
&\leq n \ln X_0 \bullet S_0 - \ln\det X_0 S_0 + \frac{\gamma^2}{2(1-\gamma)},
\end{aligned} \tag{2.20}$$

where the last relation results from applying the arithmetic-geometric mean inequality to the eigenvalues of $X_0 S_0$ (which are all real.) By (2.17) of lemma 2.1 we have

$$\Delta_1 < \left(1 - \frac{r}{n+r} - \frac{n}{n+r}\frac{\alpha}{\sqrt{n+n^2/r}}\right)\Delta_0.$$

Thus,

$$r \ln \frac{X_0 \bullet S_1}{X_0 \bullet S_0} = r \ln \frac{\Delta_1}{\Delta_0} \leq \frac{r^2}{n+r}\left(-1 + \frac{\alpha}{\sqrt{r + r^2/n}}\right). \tag{2.21}$$

Adding (2.20) and (2.21) we get

$$\psi(X_0, S_1) - \psi(X_0, S_0) \leq \frac{r^2}{n+r}\left(-1 + \frac{\alpha}{\sqrt{r+r^2/n}}\right) + \frac{\gamma^2}{2(1-\gamma)}. \tag{2.22}$$

It is easily verified that choice of $\alpha = 0.55$, $\beta = 0.3$ and $\delta = 0.1$ is consistent with all the conditions of the theorem. ∎

Based on this result we can present a projective version of the algorithm (See papar[ ] by F. Alizadeh, omitted) .