311 Final Review and Open Questions

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/ yyye

Separating and supporting hyperplane theorem

The most important theorem about the convex set is the following separating hyperplane theorem.

Theorem 1 *(Separating hyperplane theorem) Let C be a closed convex set in $\mathcal{R}^m$ and let $\mathbf{b}$ be a point exterior to C. Then there is a vector $\mathbf{y} \in \mathcal{R}^m$ such that*

$$\mathbf{b} \bullet \mathbf{y} > \sup_{x \in C} \mathbf{x} \bullet \mathbf{y}.$$

Theorem 2 *(Supporting hyperplane theorem) Let C be a closed convex set and let $\mathbf{b}$ be a point on the boundary of C. Then there is a vector $\mathbf{y} \in \mathcal{R}^m$ such that*

$$\mathbf{b} \bullet \mathbf{y} = \sup_{x \in C} \mathbf{x} \bullet \mathbf{y}.$$

Farkas' Lemma

The following results are Farkas' lemma and its variants.

Theorem 3 *Let* $A \in \mathcal{R}^{m\times n}$ *and* $\mathbf{b} \in \mathcal{R}^m$. *Then, the system* $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ *has a feasible solution* $\mathbf{x}$ *if and only if that* $-A^T\mathbf{y} \geq \mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y} > 0$ *has no feasible solution* $\mathbf{y}$.

Geometrically, Farkas' lemma means that if a vector $\mathbf{b} \in \mathcal{R}^m$ does not belong to the convex cone generated by $\mathbf{a}_{.1}, ..., \mathbf{a}_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from $\text{cone}(\mathbf{a}_{.1}, ..., \mathbf{a}_{.n})$.

Farkas' Lemma for General Cones?

Given $\mathbf{a}_i$, $i = 1, ..., m$, and $\mathbf{b} \in \mathcal{R}^m$. An analog lemma would be: the system $\{\mathbf{x}: \ \mathbf{a}_i \bullet \mathbf{x} = b_i, \ i = 1, ..., m, \ \mathbf{x} \in K\}$ has a feasible solution $\mathbf{x}$ if and only if that $-\sum_i^m y_i \mathbf{a}_i \in K^*$ and $\mathbf{b}^T \mathbf{y} > 0$ has no feasible solution $\mathbf{y}$?

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m$$

and

$$\mathcal{A}^T \mathbf{y} = \sum_i^m y_i \mathbf{a}_i.$$

Is the following a alternative system pair:

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T \mathbf{y} \in K^*, \quad \mathbf{b}^T \mathbf{y} = 1?$$

When Farkas' Lemma Holds for General Cones?

Let K be a closed and pointed convex cone in the rest of the course.

If there is $\mathbf{y}$ such that $-\mathcal{A}^T\mathbf{y} \in \text{int } K^*$, then

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1$$

are an alternative system pair.

And if there is $\mathbf{x}$ such that $\mathcal{A}^T\mathbf{x} = \mathbf{0}, \ \mathbf{x} \in \text{int } K$, then

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in K, \quad \mathbf{c} \bullet \mathbf{x} = -1 (< 0)$$

and

$$\mathbf{c} - \mathcal{A}^T\mathbf{y} \in K^*$$

are an alternative system pair.

Conic LP

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K,$$

where K is a closed and pointed convex cone.

Linear Programming (LP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = \mathcal{R}^n_+$

Second-Order Cone Programming (SOCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = SOC = \{\mathbf{x} : \ x_1 \geq \|\mathbf{x}_{2:n}\|_2\}$.

Semidefinite Programming (SDP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{S}^n$ and $K = \mathcal{S}^n_+$

p-Order Cone Programming (POCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = POC = \{\mathbf{x} : \ x_1 \geq \|\mathbf{x}_{2:n}\|_p\}$.

Dual of Conic LP

The dual problem to

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K.$$

is

$$(CLD) \quad \text{maximize} \quad \mathbf{b}^T \mathbf{y}$$
$$\text{subject to} \quad \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \in K^*,$$

where $y \in \mathcal{R}^m$, $\mathbf{s}$ is called the dual slack vector/matrix, and K^* is the dual cone of K.

Theorem 4 *(Weak duality theorem)*

$$\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T \mathbf{y} = \mathbf{x} \bullet \mathbf{s} \geq 0$$

for any feasible $\mathbf{x}$ *of (CLP) and* $(\mathbf{y}, \mathbf{s})$ *of (CLD).*

Strong Duality Theorem for CLP

Theorem 5 *The following statements hold for every pair of (CLP) and (CLD):*

i) *Let (CLP) or (CLD) be infeasible, and furthermore the other be feasible and have an interior. Then the other is unbounded.*

ii) *Let (CLP) and (CLD) be both feasible, and furthermore one of them have an interior. Then there is no duality gap between (CLP) and (CLD).*

iii) *Let (CLP) and (CLD) be both feasible and have interior. Then, both have attainable optimal solutions with no duality gap.*

In case ii), one of the optimal solution may not attainable although no gap.

General Nonlinear Optimization Problems

The question: How does one recognize an optimal solution to a nonlinearly constrained optimization problem? Let the problem have the form Consider the constrained problem again: find $\mathbf{x} \in R^n$ to

$$\begin{array}{lll} (GNO) & \inf & f(\mathbf{x}) \\ & \text{s.t.} & \mathbf{h}(\mathbf{x}) = \mathbf{0} \in R^m, \\ & & \mathbf{c}(\mathbf{x}) \leq \mathbf{0} \in R^p, \\ & & \mathbf{x} \in X \subset R^n. \end{array}$$

Lagrangian Function: $L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{h}(\mathbf{x}) + \mathbf{s}^T \mathbf{c}(\mathbf{x})$, and Lagrangian Relaxation Problem for given Lagrange multipliers $(\mathbf{y}, \mathbf{s} \geq \mathbf{0})$:

$$\begin{array}{lll} \phi(\mathbf{y}, \mathbf{s}) := & \inf & L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{h}(\mathbf{x}) + \mathbf{s}^T \mathbf{c}(\mathbf{x}) \\ & \text{s.t.} & \mathbf{x} \in X. \end{array}$$

The Lagrangian Dual Problem and Zero-Order Sufficient Condition

$$\begin{array}{lll} (LDP) & \sup & \phi(\mathbf{y},\mathbf{s}) \\ & \text{s.t.} & \mathbf{y},\ \mathbf{s} \geq \mathbf{0}. \end{array}$$

would called the Lagrangian dual of the original GNO problem:

Theorem 6 *(Dual concavity) The Lagrangian dual function $\phi(\mathbf{y},\mathbf{s})$ is a concave function.*

Theorem 7 *(Weak duality) For every $(\mathbf{y},\ \mathbf{s} \geq \mathbf{0}$, the Lagrangian dual function value $\phi(\mathbf{y},\mathbf{s})$ is less or equal to the infimum value of the original GNO problem.*

Theorem 8 *(Zero-order sufficient condition) For a feasible $\mathbf{x}$, if there is $(\mathbf{y},\ \mathbf{s} \geq \mathbf{0})$ such that $f(\mathbf{x}) = \phi(\mathbf{y},\mathbf{s})$, then $\mathbf{x}$ is a minimizer of GNO.*

Strong Duality Theorem: Zero-Oder Necessary Condition

Theorem 9 *Let (GNO) be a convex minimization problem, the infimum* f^* *of (GNO) be finite, and the suprermum of (LDP) be* ϕ^**. In addition, let (GNO) have an interior-point feasible solution, that is, there is* $\hat{\mathbf{x}}$ *such that* $\mathbf{c}(\hat{\mathbf{x}}) < \mathbf{0}$*. Then,* $f^* = \phi^*$*, and (LDP) admits a maximizer* $(\mathbf{y}^*, \mathbf{s}^* \geq \mathbf{0}$ *such that*

$$\phi(\mathbf{y}^*, \mathbf{s}^*) = f^*.$$

Furthermore, if (GNO) admits a minimizer $\mathbf{x}^*$*, then*

$$y_i^* c_i(\mathbf{x}^*) = 0, \ \forall i = 1, ..., p.$$

The assumption of “existence of an interior-point feasible solution” is usually called Constraint Qualification condition.

First-Order Necessary Condition for GNO

Theorem 10 *(First-Order or KKT necessary condition) Let* $\mathbf{x}^*$ *be a (local) minimizer of (GNO) and it is a regular point. Then, there exist multipliers* $(\mathbf{y}^*, \mathbf{s}^* \geq \mathbf{0}$ *such that*

$$\nabla_x L(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*) = \mathbf{0}$$

and $s_i^* c_i(\mathbf{x}^*) = 0, \ \forall i.$

Regular Point Qualification: For any feasible $\mathbf{x}^*$ of GNO, let active set

$$\mathcal{C}^* = \{i : \ c_i(\mathbf{x}^*) = 0, \ i = 1, ..., p\},$$

and tangent (linear) subspace

$$T^* := \{\mathbf{z} : \nabla \mathbf{h}(\mathbf{x}^*)\mathbf{z} = \mathbf{0}, \ \nabla c_i(\mathbf{x}^*)\mathbf{z} = 0 \forall i \in \mathcal{C}^*\}.$$

Then, $\mathbf{x}^*$ is a regular point if the rows of the tangent subspace matrix are linearly independent.

Second-Order Conditions for GNO

Theorem 11 *(Necessary) Let $\mathbf{x}^*$ be a (local) minimizer of (GNO) and be a regular point, and let $\mathbf{y}^*, \mathbf{s}^*$ denote Lagrange multipliers such that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$ satisfy the (first-order) KKT conditions. Then, it is necessary*

$$\mathbf{z}^T \nabla^2_{\mathbf{x}} L(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)\mathbf{z} \geq 0 \qquad \forall \mathbf{z} \in T^*.$$

Theorem 12 *(Sufficient) Let $\mathbf{x}^*$ be a regular point of (GNO) and let $\mathbf{y}^*, \mathbf{s}^*$ be the Lagrange multipliers such that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$ satisfy the (first-order) KKT conditions. Then, if in addition*

$$\mathbf{z}^T \nabla^2_{\mathbf{x}} L(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)\mathbf{z} > 0 \qquad \forall \mathbf{0} \neq \mathbf{z} \in T^*,$$

$\mathbf{x}^*$ *is a local minimizer of (GNO).*

Optimization Algorithms

Optimization algorithms tend to be iterative procedures.

Starting from a given point $\mathbf{x}^0$, they generate a sequence $\{\mathbf{x}^k\}$ of iterates (or trial solutions).

We study algorithms that produce iterates according to well determined rules–Deterministic Algorithm rather than some random selection process–Randomized Algorithm.

The rules to be followed and the procedures that can be applied depend to a large extent on the characteristics of the problem to be solved.

Classes algorithms

Depending on information of the problem used to create a new iterate:

(a) Zero-order algorithms;

(b) First-order algorithms;

(c) Second-order algorithms.

Finite versus convergent iterative methods. For some classes of optimization problems (e.g., linear and quadratic programming) there are algorithms that obtain a solution—or detect that the objective function is unbounded—in a finite number of iterations. For this reason, we call them finite algorithms.

Most algorithms encountered in nonlinear programming are not finite, but instead are convergent—or at least they are designed to be so. Their object is to generate a sequence of trial or approximate solutions that converge to a "solution."

Global Convergence Theorem

Theorem 13 *Let A be an "algorithmic mapping" defined over set X, and let sequence $\{\mathbf{x}^k\}$, starting from a given point $\mathbf{x}^0$, be generated from*

$$\mathbf{x}^{k+1} \in A(\mathbf{x}^k).$$

Let a solution set $S \subset X$ be given, and suppose

i) all points $\{\mathbf{x}^k\}$ are in a compact set;

ii) there is a continuous function $z(\mathbf{x})$ such that if $\mathbf{x} \notin S$, then $z(\mathbf{y}) < z(\mathbf{x})$ for all $\mathbf{y} \in A(\mathbf{x})$; otherwise, $z(\mathbf{y}) \le z(\mathbf{x})$ for all $\mathbf{y} \in A(\mathbf{x})$;

iii) the mapping A is closed at points outside S.

Then, the limit of any convergent subsequences of $\{\mathbf{x}^k\}$ is a solution in S.

Examples of convergence speed

The arithmetic convergence: $\left\{\frac{1}{k}\right\}$.

The q-linear convergence: $\left\{\left(\frac{1}{2}\right)^k\right\}$.

The q-quadratic convergence: $\left\{\left(\frac{1}{2}\right)^{2^k}\right\}$.

The q-superlinear convergence: $\left\{\left(\frac{1}{log(k+1)}\right)^k\right\}$.

Search direction and step-size

Typically, a nonlinear programming algorithm generates a sequence of points through an iterative scheme of the form

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{d}^k$$

where $\mathbf{d}^k$ is the search direction and α^k is the step size or step length.

The point is that once $\mathbf{x}^k$ is known, then $\mathbf{d}^k$ is some function of $\mathbf{x}^k$, and the scalar α_k may be chosen in accordance with some one-dimension -search rules.

First-Order Method: The Steepest Descent Method (SDM)

Let f be a differentiable function and we line to solve the unconstrained minimization problem $\min_{\mathbf{x}\in R^n} f(\mathbf{x})$. The solution set would be the set of stationary or KKT point of f, that is, a point $\mathbf{x}^*$ such that $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

SDM: choose $\mathbf{d}^k = -\nabla f(\mathbf{x}^k)$ as the search direction at $\mathbf{x}^k$, and select a step-size α^k.

- Optimal step-size:

$$\alpha_k = \arg\min_{\alpha} f(\mathbf{x}^k - \alpha \nabla f(\mathbf{x}^k)).$$

- Fixed step-size: $\alpha^k = \frac{1}{\beta}$ and β is β-Lipschitz coefficient of f: for any two points $\mathbf{x}$ and $\mathbf{y}$

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq \beta \|\mathbf{x} - \mathbf{y}\|.$$

- Accelerated SDM: fixed step-size with an accumulated correction.
- Barzilai and Borwein step-size: Let

$$\Delta_x^k = \mathbf{x}^k - \mathbf{x}^{k-1} \quad \text{and} \quad \Delta_g^k = \nabla f(\mathbf{x}^k) - \nabla f(\mathbf{x}^{k-1}),$$

$$\alpha^k = \frac{(\Delta_x^k)^T \Delta_g^k}{(\Delta_g^k)^T \Delta_g^k} \quad \text{or} \quad \alpha^k = \frac{(\Delta_x^k)^T \Delta_x^k}{(\Delta_x^k)^T \Delta_g^k}.$$

Convergence of SDM

The following theorem gives some conditions under which the steepest descent method will converge.

Theorem 14 *Let* $f : R^n \to R$ *be given. For a given initial point* $\mathbf{x}^0 \in R^n$*, let the level set*

$$X^0 = \{\mathbf{x} \in R^n : f(\mathbf{x}) \le f(\mathbf{x}^0)\}$$

be bounded. Assume further that f *is continuously differentiable on the convex hull of* X^0*. Let* $\{\mathbf{x}^k\}$ *be the sequence of points generated by the SDM initiated at* $\mathbf{x}^0$ *and* $f(\mathbf{x}^k)$ *be monotonously decreasing. Then every accumulation point of* $\{\mathbf{x}^k\}$ *is a stationary or KKT point of* f*.*

Second-Order Method: Newton's Method

The iteration is given by

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha^k (\nabla^2 f(\mathbf{x}^k))^{-1} \nabla f(\mathbf{x}^k).$$

Theorem 15 *Let $f(\mathbf{x})$ be twice continuously differentiable and satisfy the (second-order) β-Lipschitz condition, that is, for any two points $\mathbf{x}$ and $\mathbf{y}$*

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}) - \nabla^2 f(\mathbf{y})(\mathbf{x} - \mathbf{y})\| \leq \beta \|\mathbf{x} - \mathbf{y}\|^2$$

for a positive real number β. Also let $\mathbf{x}^*$ be a local minimizer of f at which $\nabla^2(\mathbf{x}^*)$ is positive definite. Then, provided that $\|\mathbf{x}^0 - \mathbf{x}^*\|$ is sufficiently small, the sequence generated by Newton's method converges quadratically to $\mathbf{x}^*$.

Hybrid-Order Method: the Quasi-Newton Method

In general:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha^k S^k \nabla f(\bar{\mathbf{x}})^T,$$

for a symmetric matrix S^k with a step-size scalar α^k.

SDM: $S^k = I$, α^k is decided by line search.

Newton: $S^k = (\nabla^2 f(\mathbf{x}^k))^{-1}$, $\alpha^k = 1$ or by one-dimension search.

Hibrid: $S^k = (\nabla^2 f(\mathbf{x}^k) + \lambda I)^{-1}$, $\alpha^k = 1$ or by one-dimension search.

Various methods were developed such that $S^0 = I$, and then S^k is gradually becoming $(\nabla^2 f(\mathbf{x}^k))^{-1}$....

The Augmented Lagrangian Function and Method (ALM)

We consider

$$f^* := \min \quad f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{h}(\mathbf{x}) = \mathbf{0},\ \mathbf{x} \in X.$$

Augmented Lagrangian function (ALF)

$$L_a(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{h}(\mathbf{x}) + \frac{\beta}{2} \|\mathbf{h}(\mathbf{x})\|^2.$$

ALM: Start from any $(\mathbf{x}^0 \in X, \mathbf{y}^0)$, compute a new iterate pair

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x} \in X} L_a(\mathbf{x}, \mathbf{y}^k), \text{ and } \mathbf{y}^{k+1} = \mathbf{y}^k + \beta \mathbf{h}(\mathbf{x}^{k+1}).$$

The calculation of $\mathbf{x}^{k+1}$ is used to compute the gradient vector of $\phi_a(\mathbf{y}^k)$, which is a steepest ascent direction of the dual function. Thus, the method converges just like the SDM for convex optimization, because the dual function satisfies $\frac{1}{\beta}$-Lipschitz condition. In particular, if $\mathbf{y}^k = \mathbf{y}^*$, then $\mathbf{x}^{k+1} = \mathbf{x}^*$.

The Alternating Direction Method with Multipliers (ADMM)

For the ADMM, we consider structured problem

$$\min \quad f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) \quad \text{s.t.} \quad A_1\mathbf{x}_1 + A_2\mathbf{x}_2 = \mathbf{b}.$$

Consider

$$L(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) + \mathbf{y}^T(A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}) + \frac{\beta}{2}\|A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}\|^2.$$

Then, for any given $(\mathbf{x}_1^k, \mathbf{x}_2^k, \mathbf{y}^k)$, we compute a new iterate

$$\begin{aligned}
\mathbf{x}_1^{k+1} &= \arg\min_{\mathbf{x}_1} L(\mathbf{x}_1, \mathbf{x}_2^k, \mathbf{y}^k), \\
\mathbf{x}_2^{k+1} &= \arg\min_{\mathbf{x}_2} L(\mathbf{x}_1^{k+1}, \mathbf{x}_2, \mathbf{y}^k), \\
\mathbf{y}^{k+1} &= \mathbf{y}^k + \beta(A_1\mathbf{x}_1^{k+1} + A_2\mathbf{x}_2^{k+1} - \mathbf{b}).
\end{aligned}$$

Again, we can prove that the iterates converge with the same speed for convex optimization.

Barrier Function Method for CLP

Consider the LP problem with a barrier function

$$
\begin{array}{lll}
(CLPB) & \text{minimize} & \mathbf{c}^T\mathbf{x} + \mu B_p(\mathbf{x}) \\
 & \text{s.t.} & \mathbf{x} \in \operatorname{int} \mathcal{F}_p
\end{array}
$$

and also

$$
\begin{array}{lll}
(CLDB) & \text{maximize} & \mathbf{b}^T\mathbf{y} + \mu B_d(\mathbf{s}) \\
 & \text{s.t.} & (\mathbf{y}, \mathbf{s}) \in \operatorname{int} \mathcal{F}_d,
\end{array}
$$

where μ is called the barrier (weight) parameter.

The Path-Following Algorithm

In general, one can start from an (approximate) central path point $\mathbf{x}(\mu^0)$, $(\mathbf{y}(\mu^0), \mathbf{s}(\mu^0))$, or $(\mathbf{x}(\mu^0), \mathbf{y}(\mu^0), \mathbf{s}(\mu^0))$ where μ^0 is sufficiently large.

Then, let μ^1 be a slightly smaller parameter than μ^0. Then, we compute an (approximate) central path point $\mathbf{x}(\mu^1)$, $(\mathbf{y}(\mu^1), \mathbf{s}(\mu^1))$, or $(\mathbf{x}(\mu^1), \mathbf{y}(\mu^1), \mathbf{s}(\mu^1))$. They can be updated from the previous point at μ^0 using the Newton method.

μ might be reduced at each stage by a specific factor, giving $\mu^{k+1} = \gamma\mu^k$ where γ is a fixed positive constant less than one, and k is the stage count.

This is called the primal, dual, or primal-dual path-following method.

Software Implementation

IPOPT: `https://projects.coin-or.org/Ipopt`

SEDUMI: `http://sedumi.mcmaster.ca/`

MOSEK: `http://www.mosek.com/products_mosek.html`

SDDPT3:

`http://www.math.nus.edu.sg/~mattohkc/sdpt3.html`

DSDP (Dual Semidefinite Programming Algorithm):

`http://www.stanford.edu/~yyye/Col.html`.

hsdLPsolver and more:

`http://www.stanford.edu/~yyye/matlab.html`

CVX: `http://www.stanford.edu/~boyd/cvx`