MTH 995-003: Intro to CS and Big Data Spring 2014

Lecture 4 — Jan 16, 2014

Inst. Mark Iwen *Scribe: Islam Badreldin*

Contents

1 Overview

In the last lecture, we discussed Singular Value Decomposition and its perturbation bounds, and introduced Semi-definite Programming and Convexity. In this lecture, discuss Linear Programming as a special case of Semi-definite Programming, and show examples of how to reduce other problems through algebraic manipulations into linear or semi-definite programs.

2 Linear Programming (LP)

2.1 Standard Form

Minimize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} - \mathbf{b} \geq \mathbf{0}$.

Given constants are $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^N$, and $\mathbf{A} \in \mathbb{R}^{N\times m}$. The minimization variables are $\mathbf{x} \in \mathbb{R}^m$.

2.2 Examples

Example 1. *We can re-express equality constraints in LP standard form.*

$$\begin{aligned}
&\mathbf{Ax} = \mathbf{b}\\
\Rightarrow &\mathbf{Ax} \geq \mathbf{b} \,\&\, \mathbf{Ax} \leq \mathbf{b}\\
\Rightarrow &\mathbf{Ax} \geq \mathbf{b} \,\&\, -\mathbf{Ax} \geq -\mathbf{b}\\
\Rightarrow &\begin{pmatrix} \mathbf{A} \\ -\mathbf{A} \end{pmatrix} \mathbf{x} \geq \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \end{pmatrix}
\end{aligned}$$

Equality constraints are OK for LP.

Example 2. *Compressive Sensing Recovery, aka Basis Pursuit (BP). (See Ch.3 of [FR13].)*

Minimize $||\mathbf{z}||_1$ *such that* $\mathbf{Az} = \mathbf{y}$ *(*$= \mathbf{Ax}$*, where* $\mathbf{x}$ *is the signal to be recovered), where* $\mathbf{A} \in \mathbb{R}^{m \times N}$ *and* $m \ll N$*.*

This problem can be solved as a LP by first introducing two new vectors to replace $\mathbf{z} \in \mathbb{R}^N$ *as variables. Let* $\mathbf{z}_+, \mathbf{z}_- \in \mathbb{R}^N$ *with constraints* $\mathbf{z}_+ \geq \mathbf{0}, \mathbf{z}_- \geq \mathbf{0}$ *(i.e. with only non-negative entries – we want to think of these as* $\mathbf{z} = \mathbf{z}_+ - \mathbf{z}_-$*).*

Then, we re-express the constraint as $\mathbf{A}(\mathbf{z}_+ - \mathbf{z}_-) = \mathbf{y}$*, i.e.*

$$(\mathbf{A}|-\mathbf{A}) \begin{pmatrix} \mathbf{z}_+ \\ \mathbf{z}_- \end{pmatrix} = \mathbf{y}$$

The LP problem statement is then:

Minimize $\langle (1, \ldots, 1), (\mathbf{z}_+|\mathbf{z}_-) \rangle$ *subject to*

$$\begin{aligned}
(\mathbf{A}|-\mathbf{A}) \begin{pmatrix} \mathbf{z}_+ \\ \mathbf{z}_- \end{pmatrix} &= \mathbf{y}\\
\mathbf{z}_+ \geq \mathbf{0}, \mathbf{z}_- &\geq \mathbf{0}
\end{aligned}$$

2.3 Homework Problems

Problem 5 In reference to Example 2, suppose that $\mathbf{z}^*$ has minimal $||\mathbf{z}||_1$ such that $\mathbf{Az} = \mathbf{y}$ (BP). Let ${\mathbf{z}^*}_+$ and ${\mathbf{z}^*}_-$ be the solution to the LP. Show that

$$\begin{aligned}
(z^*_+)_j > 0 \Rightarrow (z^*_-)_j = 0\\
(z^*_-)_j > 0 \Rightarrow (z^*_+)_j = 0
\end{aligned}$$

And deduce that

$$||\mathbf{z}^*||_1 = ||{\mathbf{z}^*}_+ - {\mathbf{z}^*}_-||_1$$

i.e. both BP and LP solutions have the same l_1-norm.

2.4 Relation to Semi-definite Programming (SDP)

Every LP is also a SDP, since the linear coordinate-wise inequality $\mathbf{Ax} + \mathbf{b} \geq \mathbf{0}$ can be expressed as

$$\begin{aligned}
\mathbf{Ax} + \mathbf{b} &= (\mathbf{a}_1|\mathbf{a}_2|\dots|\mathbf{a}_N)\mathbf{x} + \mathbf{b} \\
&= \mathbf{b} + \sum_{j=1}^{N} x_j \mathbf{a}_j \\
&= \begin{pmatrix} b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_N \end{pmatrix} + \sum_{j=1}^{N} x_j \begin{pmatrix} (a_j)_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (a_j)_N \end{pmatrix} \\
&= \mathbf{F}(\mathbf{x}) \geq 0
\end{aligned}$$

- SDPs are a **superset** of CS recovery algorithms, at least as far as BP goes.
- Casting problems as SDPs, or approximating solutions using SDPs, often involves re-expressing problem constraints using positive semi-definite matrices.

3 Casting Problems as SDPs

3.1 Examples

Example 3. *Having two constraints* $\mathbf{G}(\mathbf{x}) \geq 0$ *and* $\mathbf{F}(\mathbf{x}) \geq 0$ *can be re-expressed as*

$$\begin{bmatrix} \mathbf{F}(\mathbf{x}) & \mathbf{0} \\ \mathbf{0} & \mathbf{G}(\mathbf{x}) \end{bmatrix} \geq 0$$

Example 4. *Minimize the operator norm (i.e. the largest singular value) of a matrix* $\mathbf{A}(\mathbf{x}) = \sum_{j=1}^{K} x_j \mathbf{A}_j$ *over all* $\mathbf{x} \in \mathbb{R}^K$*, where* $\mathbf{A}_j \in \mathbb{R}^{p \times q}$*.*

We can cast this operator norm problem as a SDP. Introduce $t \in \mathbb{R}^+$ *as an extra variable, so we now have* $K+1$ *variables* $(t, \mathbf{x})$*. Then solve:*
Minimize t *subject to*

$$\begin{bmatrix} t\mathbf{I} & \mathbf{A}(\mathbf{x}) \\ \mathbf{A}(\mathbf{x})^{\mathrm{T}} & t\mathbf{I} \end{bmatrix} \geq 0$$

which is equivalent to

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} t\mathbf{I} & \mathbf{0} \\ \mathbf{0} & t\mathbf{I} \end{bmatrix} + \sum_{j=1}^{K} x_j \begin{bmatrix} \mathbf{0} & \mathbf{A}_j \\ \mathbf{A}_j^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \geq 0$$

Lemma 1. *The minimal* t *is the minimal largest singular value,* σ_1*, of* $\mathbf{A}(\mathbf{x})$ *over all* $\mathbf{x} \in \mathbb{R}^K$*.*

Proof: Fix $\mathbf{x} \in \mathbb{R}^K$ and let $\mathbf{A}(\mathbf{x}) = \mathbf{A} \in \mathbb{R}^{p \times q}$. Then the constraint

$$\begin{bmatrix} t\mathbf{I} & \mathbf{A}(\mathbf{x}) \\ \mathbf{A}(\mathbf{x})^{\mathrm{T}} & t\mathbf{I} \end{bmatrix} \geq 0$$

is the same as

$$[\mathbf{z}_1^{\mathrm{T}}|\mathbf{z}_2^{\mathrm{T}}]\begin{bmatrix} t\mathbf{I} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & t\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{z}_1^{\mathrm{T}} \\ \mathbf{z}_2^{\mathrm{T}}\end{bmatrix} \geq 0$$

$$\forall \mathbf{z}_1 \in \mathbb{R}^p \,\& \forall \mathbf{z}_2 \in \mathbb{R}^q$$

$$\Updownarrow$$

$$[\mathbf{z}_1^{\mathrm{T}}|\mathbf{z}_2^{\mathrm{T}}]\begin{bmatrix} t\mathbf{z}_1 + \mathbf{A}\mathbf{z}_2 \\ \mathbf{A}^{\mathrm{T}}\mathbf{z}_1 + t\mathbf{z}_2\end{bmatrix} \geq 0$$

$$\forall \mathbf{z}_1 \in \mathbb{R}^p \,\& \forall \mathbf{z}_2 \in \mathbb{R}^q$$

$$\Updownarrow$$

$$t||\mathbf{z}_1||_2^2 + \mathbf{z}_1^{\mathrm{T}}\mathbf{A}\mathbf{z}_2 + \mathbf{z}_2^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{z}_1 + t||\mathbf{z}_2||_2^2 \geq 0$$

$$\forall \mathbf{z}_1 \in \mathbb{R}^p \,\& \forall \mathbf{z}_2 \in \mathbb{R}^q$$

This last expression is minimized when we choose $\mathbf{z}_1 \& \mathbf{z}_2$ from the SVD of $\mathbf{A}$ such that

$$\mathbf{z}_2 = \mathbf{v}_1 \quad \& \mathbf{z}_1 = -\mathbf{u}_1$$

where $\mathbf{v}_1 \in \mathbb{R}^q$ is the first column of $\mathbf{V} \in \mathbb{R}^{q \times q}$, $\mathbf{u}_1 \in \mathbb{R}^p$ is the first column of $\mathbf{U} \in \mathbb{R}^{p \times p}$, and $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}}$ is the SVD of $\mathbf{A}$. The expression then becomes

$$2t - 2\sigma_1 \geq 0$$

which further reduces to $t = \sigma_1$ when minimizing t. □

3.2 Schur Complements

Suppose $\mathbf{M} \in \mathbb{S}^N$ has the block form

$$\mathbf{M} = \begin{bmatrix}\mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C}\end{bmatrix}$$

Then, the following properties must hold

i) $\mathbf{M} > 0$ iff ($\mathbf{C} > 0$ and $\mathbf{A} - \mathbf{B}\mathbf{C}^{-\mathbf{1}}\mathbf{B}^{\mathrm{T}} > 0$).

ii) $\mathbf{C} > 0 \Rightarrow$ ($\mathbf{M} \geq 0$ iff $\mathbf{A} - \mathbf{B}\mathbf{C}^{-\mathbf{1}}\mathbf{B}^{\mathrm{T}} \geq 0$).

iii) $\mathbf{A} > 0 \Rightarrow$ ($\mathbf{M} \geq 0$ iff $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-\mathbf{1}}\mathbf{B} \geq 0$).

iv) $\mathbf{M} > 0$ iff ($\mathbf{A} > 0$ and $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-\mathbf{1}}\mathbf{B} > 0$).

References

[FR13] Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing.* Birkhäuser Basel, 2013.