Linear Optimization

Daniel Mao

Contents

Chapter 1

First Chapter

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Consider the following optimization problem

$$\begin{array}{lll} \text{(P)} & \max & c^\top x \\ & \text{subject to:} & Ax \leq b. \end{array}$$

Let S denote the feasible region of (P). Let $a_1, ..., a_m \in \mathbb{R}^n$ denote the rows of A. For any $\bar{x} \in S$, define $J(\bar{x})$ by

$$J(\bar{x}) := \{i \in \{1, ..., m\} : a_i^\top x = b_i\}.$$

THEOREM 1.1. Let $A \in \mathbb{R}^{m\times n}$. Let $b \in \mathbb{R}^m$. Define $S \subseteq \mathbb{R}^n$ by $S := \{x \in \mathbb{R}^n : Ax \leq b\}$. Let $\bar{x} \in S$. Let $A^= \in \mathbb{R}^{p\times n}$ and $b^= \in \mathbb{R}^p$ be such that $A^= x = b^=$. Then $\bar{x}$ is an extreme point of S if and only if $\operatorname{rank}(A^=) = n$.

THEOREM 1.2. Let $\bar{x} \in S$. Then $\bar{x}$ is optimal for (P) if and only if $c \in \operatorname{cone}(\{a_i\}_{i\in J})$.

Proof. ($\Rightarrow$) Suppose that $\bar{x}$ is an optimal solution to (P). Then by Strong Duality, the dual (D) has an optimal solution $\bar{y}$. Then Complementary Slackness holds. So $\forall i \in \{1, ..., m\} \setminus J$, we have $\bar{y}_i = 0$; and $\forall i \in J$, $a_i^\top x = b_i$. Since $\bar{y}$ is an optimal solution to (D), we have $A^\top \bar{y} = \sum_{i=1}^m \bar{y}_i a_i = c$ and $\bar{y} \geq 0$. So

$$c = \sum_{i=1}^m \bar{y}_i a_i = \sum_{i\in J} \bar{y}_i a_i + \sum_{i\notin J} \underbrace{\bar{y}_i}_{=0} a_i = \sum_{i\in J} \bar{y}_i a_i.$$

So $c \in \operatorname{cone}(\{a_i\}_{i \in J})$.

($\Leftarrow$) Suppose that $c \in \operatorname{cone}(\{a_i\}_{i \in J})$. Then $\exists \bar{y} \in \mathbb{R}^m_+$ such that $c = \sum_{i \in J} \bar{y}_i a_i$ and $\forall i \notin J$, $\bar{y}_i = 0$. So

$$c = \sum_{i \in J} \bar{y}_i a_i = \sum_{i \in J} \bar{y}_i a_i + \sum_{i \notin J} \underbrace{\bar{y}_i}_{=0} a_i = A^\top \bar{y}.$$

So $\bar{y}$ is a feasible solution to (D). Notice that by construction of $\bar{y}$, Complementary Slackness holds for $\bar{x}$ and $\bar{y}$. So $\bar{x}$ is an optimal solution to (P) and $\bar{y}$ is an optimal solution to (D). $\square$

Chapter 2

Interior-Point Methods

Contents

2.1 Primal-Dual Methods

Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Assume that $\operatorname{rank}(A) = m$. Consider the linear optimization problem and its dual:

$$
\begin{array}{llllll}
\text{(LP)} & \displaystyle\min_{x} & c^\top x & \text{(LD)} & \displaystyle\max_{\lambda,s} & b^\top \lambda \\
& \text{subject to:} & x \in \mathbb{R}^n, & & \text{subjec to:} & \lambda \in \mathbb{R}^m, s \in \mathbb{R}^n, \\
& & Ax = b, & & & A^\top \lambda + s = c, \\
& & x \geq \mathbb{0}_n & & & s \geq \mathbb{0}_m.
\end{array}
$$

PROPOSITION 2.1. The solutions are characterized by the KKT system

$$
\begin{cases}
A^\top \lambda + s = c, \\
Ax = b, \\
x_i s_i = 0, \quad \forall i \in \{1, ..., n\}, \\
x \geq \mathbb{0}_n, s \geq \mathbb{0}_n.
\end{cases}
$$

Algorithm 1: Primal-Dual Path-Following

Input: $(x^{(0)}, \lambda^{(0)}, s^{(0)})$ with $(x^{(0)}, s^{(0)}) > \mathbb{0}_n$.

for $k \in \mathbb{Z}_+$ **do**

Choose $\sigma_k \in [0, 1]$;

Solve the system

$$
\begin{bmatrix} 0 & A^\top & I \\ A & \mathbb{0}_{m\times m} & \mathbb{0}_{m\times m} \\ S^{(k)} & \mathbb{0}_{m\times m} & X^{(k)} \end{bmatrix}
\begin{bmatrix} \Delta x^{(k)} \\ \Delta \lambda^{(k)} \\ \Delta s^{(k)} \end{bmatrix}
=
\begin{bmatrix} -A^\top \lambda^{(k)} - s^{(k)} + c \\ -Ax^{(k)} + b \\ -X^{(k)} S^{(k)} e + \sigma_k \mu_k e \end{bmatrix}
$$

where $X^{(k)} := \operatorname{Diag}(x^{(k)})$, $S^{(k)} := \operatorname{Diag}(s^{(k)})$, and $\mu_k := \frac{1}{n} x^{(k)\top} s^{(k)}$;

Choose α_k such that $(x^{(k+1)}, s^{(k+1)}) > \mathbb{0}_n$;

Set $(x^{(k+1)}, \lambda^{(k+1)}, s^{(k+1)}) := (x^{(k)}, \lambda^{(k)}, s^{(k)}) + \alpha_k(\Delta x^{(k)}, \Delta \lambda^{(k)}, \Delta s^{(k)})$;

DEFINITION 2.2 (Primal-Dual Feasible Set/Strictly Feasible Set)**.** We define the **feasible set** $\mathcal{F}$ and **strictly feasible set** $\mathcal{F}^o$ to be

$$
\mathcal{F} := \{(x, \lambda, s) : Ax = b, A^\top \lambda + s = c, (x, s) \geq \mathbb{0}_n\},
$$

$$\mathcal{F}^o := \{(x, \lambda, s) : Ax = b, A^\top \lambda + s = c, (x, s) > \mathbb{0}_n\}.$$

DEFINITION 2.3 (Central Path)**.** We define the **central path** $\mathcal{C}$ to be

$$\mathcal{C} := \{(x_\tau, \lambda_\tau, s_\tau) : \tau > 0\}$$

where for each $\tau > 0$, $(x_\tau, \lambda_\tau, s_\tau)$ is a solution to the following system

$$\begin{cases} A^\top \lambda + s = c, \\ Ax = b, \\ x_i s_i = \tau, \quad \forall i \in \{1, ..., n\}, \\ (x, s) > \mathbb{0}_n. \end{cases}$$

PROPOSITION 2.4. The above system is also the optimality conditions for a logarithmic-barrier formulation

$$\begin{array}{lll} \text{(P)} & \min\limits_{x} & c^\top x - \tau \sum_{i=1}^{n} \ln(x_i) \\ & \text{subject to:} & x \in \mathbb{R}^n, \\ & & Ax = b, \\ & & x > \mathbb{0}_n. \end{array}$$

Proof. Form the Lagrangian function $\mathcal{L} : \mathbb{R}^n \oplus \mathbb{R}^m \to \mathbb{R}$ as

$$\mathcal{L}(x, \lambda) = f(x) - \lambda^\top (Ax - b).$$

Then

$$\begin{aligned} \nabla_x \mathcal{L}(x, \lambda) &= c - \left[\frac{\tau}{x_i}\right]_{i=1}^{n} - A^\top \lambda \text{ and} \\ \nabla_\lambda \mathcal{L}(x, \lambda) &= Ax - b. \end{aligned}$$

The KKT conditions for this problem is

$$\begin{cases} c_i - \dfrac{\tau}{x_i} - a_i^\top \lambda = 0, \quad \forall i \in \{1, ..., n\}, \\ Ax - b = \mathbb{0}_n, \\ \lambda^\top (Ax - b) = \mathbb{0}_m. \end{cases}$$

□

PROPOSITION 2.5. The solutions $(x_\tau, \lambda_\tau, s_\tau)$ are unique if and only if $\mathcal{F}^o \neq \varnothing$.

DEFINITION 2.6 (Neighborhoods of the Central Path)**.** We define the neighborhoods $\mathcal{N}_2$ and $\mathcal{N}_{-\infty} : (0,1] \to \mathcal{P}(\mathcal{F}^o)$ by

$$\mathcal{N}_2(\theta) := \{(x, \lambda, s) \in \mathcal{F}^o : \|XSe - \mu e\|_2 \leq \theta\mu\},$$
$$\mathcal{N}_{-\infty}(\gamma) := \{(x, \lambda, s) \in \mathcal{F}^o : x_i s_i \geq \gamma\mu, \quad \forall i \in \{1, ..., n\}\}$$

where $\mu := \dfrac{1}{n} x^\top s$ is the duality measure.

Algorithm 2: Long-Step Path-Following

Input: $\gamma, \sigma_{\min}, \sigma_{\max}$ with $\gamma \in (0,1)$ and $0 < \sigma_{\min} \leq \sigma_{\max} < 1$ and $(x^{(0)}, \lambda^{(0)}, s^{(0)}) \in \mathcal{N}_{-\infty}(\gamma)$

for $k \in \mathbb{Z}_+$ **do**

- Choose $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$;
- Solve the system

$$\begin{bmatrix} 0 & A^\top & I \\ A & \mathbb{0}_{m\times m} & \mathbb{0}_{m\times m} \\ S^{(k)} & \mathbb{0}_{m\times m} & X^{(k)} \end{bmatrix} \begin{bmatrix} \Delta x^{(k)} \\ \Delta\lambda^{(k)} \\ \Delta s^{(k)} \end{bmatrix} = \begin{bmatrix} -A^\top \lambda^{(k)} - s^{(k)} + c \\ -Ax^{(k)} + b \\ -X^{(k)} S^{(k)} e + \sigma_k \mu_k e \end{bmatrix}$$

  where $X^{(k)} := \mathrm{Diag}(x^{(k)})$, $S^{(k)} := \mathrm{Diag}(s^{(k)})$, and $\mu_k := \dfrac{1}{n} x^{(k)\top} s^{(k)}$;
- Choose $\alpha_k \in [0,1]$ largest such that $(x^{(k)}(\alpha), \lambda^{(k)}(\alpha), s^{(k)}(\alpha)) \in \mathcal{N}_{-\infty}(\gamma)$;
- Set $(x^{(k+1)}, \lambda^{(k+1)}, s^{(k+1)}) := (x^{(k)}, \lambda^{(k)}, s^{(k)}) + \alpha_k(\Delta x^{(k)}, \Delta\lambda^{(k)}, \Delta s^{(k)})$;

2.2 Convergence Analysis

LEMMA 2.7. Let $u, v \in \mathbb{R}^n$ with $u^\top v \geq 0$. Then

$$\| \operatorname{Diag}(u) \operatorname{Diag}(v)e\|_2 \leq 2^{-3/2}\|u+v\|_2^2.$$

Proof.

$$\begin{aligned}
\| \operatorname{Diag}(u) \operatorname{Diag}(v)e\|_2 &= \left[\sum_{i=1}^{n}(u_iv_i)^2\right]^{1/2} = \left[\sum_{i\in\mathcal{P}}(u_iv_i)^2 + \sum_{i\in\mathcal{M}}(u_iv_i)^2\right]^{1/2} \\
&\leq \left[(\sum_{i\in\mathcal{P}}|u_iv_i|)^2 + (\sum_{i\in\mathcal{M}}|u_iv_i|)^2\right]^{1/2} \\
&\leq \left[2(\sum_{i\in\mathcal{P}}|u_iv_i|)^2\right]^{1/2} = \sqrt{2}\sum_{i\in\mathcal{P}}|u_iv_i| \\
&\leq \sqrt{2}\sum_{i\in\mathcal{P}}\left[\frac{1}{4}(u_i+v_i)^2\right] = 2^{-3/2}\sum_{i\in\mathcal{P}}(u_i+v_i)^2 \\
&\leq 2^{-3/2}\sum_{i=1}^{n}(u_i+v_i)^2 = 2^{-3/2}\|u+v\|_2^2.
\end{aligned}$$

□

LEMMA 2.8. If $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma)$, then

$$\|\Delta X \Delta S e\|_2 \leq 2^{-3/2}(1+1/\gamma)n\mu.$$

THEOREM 2.9. Let $(\mu_k)_{k\in\mathbb{Z}_{++}}$ be the sequence of duality measures generated by Algorithm 2. Then there is some δ that depends only on $\gamma, \sigma_{\min}, \sigma_{\max}$ such that

$$\mu_{k+1} \leq (1-\frac{\delta}{n})\mu_k.$$

Proof. I first claim that

$$\forall\alpha \in \left[0, 2^{3/2}\frac{\sigma_k}{n}\gamma\frac{1-\gamma}{1+\gamma}\right], \quad (x^{(k)}, \lambda^{(k)}, s^{(k)}) + \alpha(\Delta x^{(k)}, \Delta\lambda^{(k)}, \Delta s^{(k)}) \in \mathcal{N}_{-\infty}(\gamma).$$

Let $\alpha \in \left[0, 2^{3/2}\frac{\sigma_k}{n}\gamma\frac{1-\gamma}{1+\gamma}\right]$ be arbitrary. Observe that

$$\mu_k(\alpha) = \frac{1}{n}[x^{(k)}(\alpha)]^\top[s^{(k)}(\alpha)]$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} x_i^{(k)} s_i^{(k)}(1-\alpha) + \frac{1}{n}\sum_{i=1}^{n}\alpha\sigma_k\mu_k + \frac{1}{n}\sum_{i=1}^{n}\alpha\Delta x_i^{(k)}\Delta s_i^{(k)} \\
&= (1-\alpha)\mu_k + \alpha\sigma_k\mu_k + \frac{\alpha^2}{n}\underbrace{[\Delta x^{(k)}]^\top[\Delta s^{(k)}]}_{=0} \\
&= (1-\alpha+\alpha\sigma_k)\mu_k.
\end{aligned}
$$

Then $\forall i \in \{1, ..., n\}$,

$$
\begin{aligned}
&x_i^{(k)}(\alpha)s_i^{(k)}(\alpha) = (x_i^{(k)} + \alpha\Delta x_i^{(k)})(s_i^{(k)} + \alpha\Delta s_i^{(k)}) \\
&= x_i^{(k)}s_i^{(k)} + \alpha x_i^{(k)}\Delta s_i^{(k)} + \alpha s_i^{(k)}\Delta x_i^{(k)} + \alpha^2\Delta x_i^{(k)}\Delta s_i^{(k)} \\
&= x_i^{(k)}s_i^{(k)} + \alpha\Big[-x_i^{(k)}s_i^{(k)} + \sigma_k\mu_k\Big] + \alpha^2\Delta x_i^{(k)}\Delta s_i^{(k)},\ \text{since } \Delta x^{(k)} \text{ and } \Delta s^{(k)} \text{ solves the system} \\
&= x_i^{(k)}s_i^{(k)}(1-\alpha) + \alpha\sigma_k\mu_k + \alpha^2\Delta x_i^{(k)}\Delta s_i^{(k)} \\
&\geq x_i^{(k)}s_i^{(k)}(1-\alpha) + \alpha\sigma_k\mu_k - \alpha^2|\Delta x_i^{(k)}\Delta s_i^{(k)}| \\
&\geq \gamma\mu_k(1-\alpha) + \alpha\sigma_k\mu_k - \alpha^2|\Delta x_i^{(k)}\Delta s_i^{(k)}|,\ \text{since } x^{(k)}s^{(k)} \in \mathcal{N}_{-\infty} \\
&\geq \gamma\mu_k(1-\alpha) + \alpha\sigma_k\mu_k - \alpha^2 2^{-3/2}(1+1/\gamma)n\mu_k,\ \text{by Lemma} \\
&\geq \gamma(1-\alpha+\alpha\sigma_k)\mu_k,\ \text{since } \alpha \leq 2^{3/2}\frac{\sigma_k}{n}\gamma\frac{1-\gamma}{1+\gamma} \\
&= \gamma\mu_k(\alpha),\ \text{by above.}
\end{aligned}
$$

That is, $x_i^{(k)}(\alpha)s_i^{(k)}(\alpha) \geq \gamma\mu_k(\alpha)$. So $(x^{(k)}(\alpha), \lambda^{(k)}(\alpha), s^{(k)}(\alpha)) \in \mathcal{N}_{-\infty}(\gamma)$. By definition of the algorithm, we have

$$\alpha_k \geq 2^{3/2}\frac{\sigma_k}{n}\gamma\frac{1-\gamma}{1+\gamma}.$$

Now

$$\mu_{k+1} = (1-\alpha_k(1-\sigma_k))\mu_k \leq (1 - 2^{3/2}\frac{\sigma_k}{n}\gamma\frac{1-\gamma}{1+\gamma}(1-\sigma_k))\mu_k.$$

Define

$$\delta := 2^{3/2}\gamma\frac{1-\gamma}{1+\gamma}\min\Big\{\sigma_{\min}(1-\sigma_{\min}), \sigma_{\max}(1-\sigma_{\max})\Big\}.$$

Then we have

$$\mu_{k+1} \leq (1 - \frac{\delta}{n})\mu_k.$$

$\square$

COROLLARY 2.10. Let $\varepsilon \in (0,1)$ and $\gamma \in (0,1)$. Suppose that $(x^{(0)}, \lambda^{(0)}, s^{(0)}) \in \mathcal{N}_{-\infty}$. Then there is an index K with $K \in O(n \log(1/\varepsilon))$ such that

$$\mu_k \leq \varepsilon \mu_0, \quad \forall k \geq K.$$