Tutorial on Robust Interior Point

May 5, 2021

In this note, we give a short self-contained proof for interior point method and its robust version. Consider the primal linear program

$$\min_{\mathbf{A}x=b,x\in\mathbb{R}^n_{\geq 0}} c^\top x \tag{P}$$

and its dual

$$\max_{\mathbf{A}^\top y+s=c,s\in\mathbb{R}^n_{\geq 0}} b^\top y. \tag{D}$$

where $\mathbf{A}\in\mathbb{R}^{d\times n}$ and $\mathbb{R}_{\geq}=\{x\geq 0\}$. The feasible region for both programs are $\mathcal{P}=\{x\in\mathbb{R}^n_{\geq}:\mathbf{A}x=b\}$ and $\mathcal{D}=\{s\in\mathbb{R}^n_{\geq}:\mathbf{A}^\top y+s=c \text{ for some } y\}$. We define their interior $\mathcal{P}^\circ=\{x\in\mathbb{R}^n_{>}:\mathbf{A}x=b\}$ and $\mathcal{D}^\circ=\{s\in\mathbb{R}^n_{>}:\mathbf{A}^\top y+s=c \text{ for some } y\}$.

To motivate interior point methods, we recall the optimality condition for linear programs:

Theorem 1 (Complementary Slackness)**.** *For any $x\in\mathcal{P}$ and $s\in\mathcal{D}$, x and s are optimal if and only if $x^\top s=0$. Furthermore, if both $\mathcal{P}$ and $\mathcal{D}$ are non-empty, there are optimal $x^*\in\mathcal{P}$ and $s^*\in\mathcal{D}$ such that $(x^*)^\top s^*=0$ and $x^*+s^*>0$.*

More generally, the quantity $x^\top s$ measures the duality gap of the feasible solution:

Lemma 2 (Duality Gap)**.** *For any $x\in\mathcal{P}$ and $s\in\mathcal{D}$, the duality gap $c^\top x-b^\top y=x^\top s$. In particular $c^\top x\leq \min_{x\in\mathcal{P}} c^\top x+x^\top s$.*

Proof. Using $\mathbf{A}x=b$ and $\mathbf{A}^\top y+s=c$, we can compute the duality gap as follows

$$c^\top x-b^\top y=c^\top x-(\mathbf{A}x)^\top y=c^\top x-x^\top(\mathbf{A}y)=x^\top s.$$

By weak duality, we have

$$c^\top x=b^\top y+x^\top s\leq \max_{\mathbf{A}^\top y+s=c,s\in\mathbb{R}^n_+} b^\top y+x^\top s\leq \min_{x\in\mathcal{P}} c^\top x+x^\top s.$$

□

The main implication of Lemma 2 is that any feasible (x,s) with small $x^\top s$ is a nearly optimal solution of the linear program. This leads us to the primal-dual algorithms in which we start with a feasible primal and dual solution (x,s) and iteratively update the solution to decrease the duality gap $x^\top s$.

1 Interior Point Method

In this note, we discuss the short-step interior point method. The algorithm maintains a pair $(x,s)\in\mathcal{P}^\circ\times\mathcal{D}^\circ$ and a scalar $t>0$ satisfying the invariant $\|\frac{xs}{t}-1\|_2\leq\frac{1}{4}$. Each step, it decreases t by a factor of $1-\Omega(n^{-1/2})$ while maintaining the invariant.

1.1 Basic Property of a Step

To see why there is a pair $(x,s)\in\mathcal{P}^\circ\times\mathcal{D}^\circ$ satisfying the invariant, we prove the following generalization:

Lemma 3 (Quadrant Representation of Primal-Dual)**.** *Suppose $\mathcal{P}$ is non-empty and bounded. For any positive vector $\mu\in\mathbb{R}^n_{>}$, there is an unique pair $(x_\mu,s_\mu)\in\mathcal{P}^\circ\times\mathcal{D}^\circ$ such that $x_\mu s_\mu=\mu$.*

Proof. Fix $\mu \in \mathbb{R}^n_>$. We define the function

$$f_\mu(x) = c^\top x - \sum_{i=1}^n \mu_i \ln x_i$$

and let $x_\mu = \arg\min_{x\in\mathcal{P}} f_\mu(x)$. Since $\mathcal{P}$ is non-empty and bounded and since f_μ is strictly convex, such unique x exists. Furthermore, since $f_\mu(x) \to +\infty$ as $x_i \to 0$ for any i, we have that $x_\mu \in \mathcal{P}^\circ$.

By the optimality condition for f_μ, there is a vector y such that

$$c - \frac{\mu}{x} = \mathbf{A}^\top y.$$

Define $s_\mu = \frac{\mu}{x_\mu}$, then one can check that $s_\mu \in \mathcal{D}$ and $x_\mu s_\mu = \mu$.

For the uniqueness, if $(x, s) \in \mathcal{P}^\circ \times \mathcal{D}^\circ$ and $xs = \mu$, then x satisfies the optimality condition for f_μ. Since f_μ is strictly convex, such x is unique. $\square$

Lemma 3 shows that any point in $\mathcal{P}^\circ \times \mathcal{D}^\circ$ is uniquely represented by a positive vector μ. Interior point methods move μ uniform to 0 while maintaining its corresponding x_μ. Now, we discuss how to find (x_μ, s_μ) given a nearby interior feasible point (x, s). Namely, how to move (x, s) to $(x + \delta_x, s + \delta_s)$ such that it satisfies the equation

$$\begin{aligned}
(x+\delta_x)(s+\delta_s) &= \mu,\\
\mathbf{A}(x+\delta_x) &= b,\\
\mathbf{A}^\top(y+\delta_y) + (s+\delta_s) &= c,\\
(x+\delta_x, s+\delta_s) &\in \mathbb{R}^{2n}_{>0}.
\end{aligned}$$

Although the equation above involves y, our approximate solution does not need to know y. By ignoring the second order term $\delta_x\delta_s$ on the equation above and the inequality constraint, we can simplify the formula a little bit by using $\mathbf{A}x = 0$ and $\mathbf{A}^\top y + s = c$:

$$\begin{aligned}
xs + \mathbf{S}\delta_x + \mathbf{X}\delta_s &= \mu, \qquad (1.1)\\
\mathbf{A}\delta_x &= 0,\\
\mathbf{A}^\top\delta_y + \delta_s &= 0,
\end{aligned}$$

where $\mathbf{X}$ and $\mathbf{S}$ are the diagonal matrix with diagonal x and s. In the following Lemma, we show how to write the step above using a projection matrix.

Lemma 4. *Suppose that $\mathbf{A}$ has full row rank and $(x, s) \in \mathcal{P}^\circ \times \mathcal{D}^\circ$. Then, the unique solution for the linear system (1.1) is given by*

$$\begin{aligned}
\mathbf{X}^{-1}\delta_x &= (\mathbf{I} - \mathbf{P})(\delta_\mu/\mu),\\
\mathbf{S}^{-1}\delta_s &= \mathbf{P}(\delta_\mu/\mu)
\end{aligned}$$

where $\delta_\mu = \mu - xs$ and $\mathbf{P} = \mathbf{S}^{-1}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{X}$.

Proof. Note that the step satisfies $\mathbf{S}\delta_x + \mathbf{X}\delta_s = \delta_\mu$. Multiply both sides by $\mathbf{A}\mathbf{S}^{-1}$ and using $\mathbf{A}\delta_x = 0$, we have

$$\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\delta_s = \mathbf{A}\mathbf{S}^{-1}\delta_\mu.$$

Now, we use that $\mathbf{A}^\top\delta_y + \delta_s = 0$ and get

$$\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top\delta_y = -\mathbf{A}\mathbf{S}^{-1}\delta_\mu.$$

Since $\mathbf{A} \in \mathbb{R}^{m\times n}$ has full row rank and $\mathbf{S}^{-1}\mathbf{X}$ is invertible, we have that $\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top$ is invertible and $\delta_y = -(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{S}^{-1}\delta_\mu$ and

$$\delta_s = \mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{S}^{-1}\delta_\mu.$$

Putting it into $\mathbf{S}\delta_x + \mathbf{X}\delta_s = \delta_\mu$, we have

$$\delta_x = \mathbf{S}^{-1}\delta_\mu - \mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{S}^{-1}\delta_\mu.$$

The result follows from the definition of $\mathbf{P}$. $\square$

1.2 Lower Bounding Step Size

The efficiency of interior point methods depends on how large the step we can take while stays within the domain. We first study the step operator $(\mathbf{I}-\mathbf{P})$ and $\mathbf{P}$. The following Lemma shows that $\mathbf{P}$ is a nearly-orthogonal projection matrix when μ is close to a constant vector. Hence, the relative changes of $\mathbf{X}^{-1}\delta_x$ and $\mathbf{S}^{-1}\delta_s$ are essentially the orthogonal decomposition of the relative step δ_μ/μ on μ.

Lemma 5. *Under the assumption in Lemma 4, $\mathbf{P}$ is a projection matrix such that $\|\mathbf{P}v\|_\mu \leq \|v\|_\mu$ for any $v \in \mathbb{R}^n$. Similarly, we have that $\|(\mathbf{I}-\mathbf{P})v\|_\mu \leq \|v\|_\mu$.*

Proof. $\mathbf{P}$ is a projection because $\mathbf{P}^2 = \mathbf{P}$. Define the orthogonal projection

$$\mathbf{P}_{\text{orth}} = \mathbf{S}^{-1/2}\mathbf{X}^{1/2}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{X}^{1/2}\mathbf{S}^{-1/2},$$

then we have

$$\begin{aligned}
\|\mathbf{P}v\|_\mu^2 &= v^\top\mathbf{X}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{S}\mathbf{S}^{-1}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{X}v \\
&= v^\top\mathbf{S}^{1/2}\mathbf{X}^{1/2}\mathbf{P}_{\text{orth}}\mathbf{S}^{1/2}\mathbf{X}^{1/2}v \\
&\leq v^\top\mathbf{S}^{1/2}\mathbf{X}^{1/2}\mathbf{S}^{1/2}\mathbf{X}^{1/2}v = \|v\|_\mu^2.
\end{aligned}$$

The calculation for $\|(\mathbf{I}-\mathbf{P})v\|_\mu$ is similar. $\square$

Now, we give a lower bound of the largest feasible step size:

Lemma 6. *We have that $\|\mathbf{X}^{-1}\delta_x\|_\infty^2 \leq \frac{1}{\min_i \mu_i}\|\delta_\mu/\mu\|_\mu^2$ and $\|\mathbf{S}^{-1}\delta_s\|_\infty^2 \leq \frac{1}{\min_i \mu_i}\|\delta_\mu/\mu\|_\mu^2$. In particular, if $\|\delta_\mu/\mu\|_\mu^2 < \min_i \mu_i$, we have $(x+\delta_x, s+\delta_s) \in \mathcal{P}^\circ \times \mathcal{D}^\circ$*

Proof. For $\|\mathbf{X}^{-1}\delta_x\|_\infty$, we have $\min_i \mu_i\|\mathbf{X}^{-1}\delta_x\|_\infty^2 \leq \|\mathbf{X}^{-1}\delta_x\|_\mu^2$ and hence

$$\|\mathbf{X}^{-1}\delta_x\|_\infty^2 \leq \frac{1}{\min_i \mu_i}\|\mathbf{X}^{-1}\delta_x\|_\mu^2 = \frac{1}{\min_i \mu_i}\|(\mathbf{I}-\mathbf{P})(\delta_\mu/\mu)\|_\mu^2 \leq \frac{1}{\min_i \mu_i}\|\delta_\mu/\mu\|_\mu^2.$$

The proof for $\|\mathbf{S}^{-1}\delta_s\|_\infty$ is similar.

Hence, if $\|\delta_\mu/\mu\|_\mu^2 < \min_i \mu_i$, we have that $\|\mathbf{X}^{-1}\delta_x\|_\infty < 1$ and $\|\mathbf{S}^{-1}\delta_s\|_\infty < 1$. Therefore, $x+\delta_x$ and $s+\delta_s$ are feasible. $\square$

To decrease μ uniformly, we set $\delta_\mu = -h\mu$ for some step size h. To ensure the feasibility, we need $\|\delta_\mu/\mu\|_\mu^2 \leq \min_i \mu_i$ and this gives the maximum step size

$$h = \sqrt{\frac{\min_i \mu_i}{\sum_i \mu_i}}. \tag{1.2}$$

Note that the above quantity maximizes at $n^{-1/2}$ when μ is a constant vector.

1.3 Staying within ℓ_2 distance

Since the step size (1.2) maximizes when μ is a constant vector. A natural approach is to control μ vector ℓ_2 close to a constant t. This motivates the following algorithm:

Algorithm 1: `L2InteriorPoint`$(\mathbf{A}, x, s, t, \epsilon)$

Invariant: $(x,s) \in \mathcal{P}^\circ \times \mathcal{D}^\circ$ and $\|xs - t\|_2 \leq \frac{t}{3}$.
Let $h = 1/(9\sqrt{n})$ where n is the number of columns in $\mathbf{A}$.
while $t > \epsilon/(2n)$ **do**
  Let $\mu = xs$ and $\delta_\mu = (1-h)t - \mu$.
  Let $\mathbf{P} = \mathbf{S}^{-1}\mathbf{A}^\top(\mathbf{A}\mathbf{S}^{-1}\mathbf{X}\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{X}$.
  Let $\delta_x = \mathbf{X}(\mathbf{I}-\mathbf{P})(\delta_\mu/\mu)$ and $\delta_s = \mathbf{S}\mathbf{P}(\delta_\mu/\mu)$.
  Set $x \leftarrow x+\delta_x$, $s \leftarrow s+\delta_s$ and $t \leftarrow (1-h)t$.
end
Return (x,s)

Note that the algorithm requires some initial point (x,s) with $xs \approx t$ and we will show how to modify the linear programs such that the optimal point is approximately the same while $x = 1_n$, $s = 1_n$ are feasible primal dual point.

Theorem 7. *Suppose that the input satisfies $(x,s) \in \mathcal{P}^\circ \times \mathcal{D}^\circ$ and $\|xs - t\|_2 \leq \frac{t}{3}$. Then it holds throughout the algorithm* `L2InteriorPoint`*. Furthermore, the output satisfies $x^\top s \leq \epsilon$.*

Proof. We prove by induction that $\|xs - t\|_2 \leq \frac{t}{3}$. It holds for the first step by the assumption.

Let $x' = x + \delta_x$, $s' = s + \delta_s$ and $t' = (1-h)t$. Note that

$$\begin{aligned} x's' - t' =& (x+\delta_x)(s+\delta_s) - (1-h)t \\ =& \mu + \mathbf{S}\delta_x + \mathbf{X}\delta_s + \delta_x\delta_s - (1-h)t. \end{aligned}$$

Lemma 4 shows that $\mathbf{S}\delta_x + \mathbf{X}\delta_s = (1-h)t - \mu$. Hence, we have

$$x's' - t' = \delta_x\delta_s = \mathbf{X}^{-1}\delta_x \cdot \mathbf{S}^{-1}\delta_s \cdot \mu$$

Using this, we have

$$\begin{aligned} \|x's' - t'\|_2 &\leq \|\mu^{1/2}\mathbf{X}^{-1}\delta_x\|_2 \|\mu^{1/2}\mathbf{S}^{-1}\delta_s\|_2 \\ &= \|\mathbf{X}^{-1}\delta_x\|_\mu \|\mathbf{S}^{-1}\delta_s\|_\mu \\ &\leq \|\delta_\mu/\mu\|_\mu^2 \end{aligned}$$

where we used Lemma 5 at the end.

For the last term, we note that

$$\begin{aligned} \|\delta_\mu/\mu\|_\mu &= \|(1-h)(t-\mu)/\mu + h\|_\mu \\ &\leq \|xs - t\|_{\mu^{-1}} + h\|1\|_\mu. \end{aligned}$$

Since $\|\mu - t\|_2 \leq \frac{t}{3}$, we have $\min_i \mu_i \geq \frac{2t}{3}$ and $\max_i \mu_i \leq \frac{4}{3}t$. Using $h = \sqrt{n}/9$, we have

$$\|\delta_\mu/\mu\|_\mu \leq \sqrt{\frac{3}{2t}}\|xs - t\|_2 + h\sqrt{\frac{4}{3}n} \leq \sqrt{\frac{t}{6}} + h\sqrt{\frac{4}{3}n} \leq 0.537\sqrt{t}.$$

Hence, we have $\|x's' - t'\|_2 \leq \|\delta_\mu/\mu\|_\mu^2 \leq 0.289t \leq t'/3$. Furthermore, $\|\delta_\mu/\mu\|_\mu^2 < \min_i \mu_i$ which implies (x,s) is feasible (Lemma 6).

For the last iteration, we have $t \leq \epsilon/(2n)$ and hence $x^\top s \leq \sum_i \mu_i \leq \frac{4}{3}nt \leq \epsilon$. □

1.4 Finding the Initial Point

To find the initial point, we handle it by extending the problem to slightly higher dimension. The following lemma shows that we can insert extra blocks in $\overline{\mathbf{A}}$ to force $(x,s) = (1,1,\cdots,1)$ on the central path. The proof is a gadget-style proof and does not give intuition on interior point methods. Hence, we defer it to appendix.

Lemma 8. *Consider a linear program $\min_{\mathbf{A}x=b,x\geq 0} c^\top x$ with n variables and m constraints. Assume that*

1. *Diameter: For any $x \geq 0$ with $\mathbf{A}x = b$, we have that $\|x\|_\infty \leq R$.*
2. *Lipschitz constant of the objective: $\|c\|_\infty \leq L$.*

For any $0 < \delta \leq \frac{1}{2}$, the modified linear program $\min_{\overline{\mathbf{A}}\overline{x}=\overline{b},\overline{x}\geq 0} \overline{c}^\top \overline{x}$ with

$$\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{D} & 0_d \\ u^\top & 0_d^\top & 1 \end{bmatrix}, \overline{b} = \begin{bmatrix} \frac{1}{R}b \\ U+1 \end{bmatrix}, \text{ and } \overline{c} = \begin{bmatrix} \delta/L \cdot c \\ 1_d \\ 0 \end{bmatrix}$$

with $u = 1_n - \delta/L \cdot c$, $\mathbf{D} = \mathrm{diag}(\frac{1}{R}b - \mathbf{A}1_n)$ and $U = \sum_{i\in[n]} u_i$ satisfies the following:

1. $\overline{x} = \begin{bmatrix} 1_n \\ 1_d \\ 1 \end{bmatrix}, \overline{y} = \begin{bmatrix} 0_d \\ -1 \end{bmatrix}$ *and* $\overline{s} = \begin{bmatrix} 1_n \\ 1_d \\ 1 \end{bmatrix}$ *are feasible primal dual vectors.*

2. *For any feasible primal dual vectors* $(\overline{x}, \overline{s})$ *with duality gap at most* δ^2*, the vector* $\hat{x} = R \cdot \overline{x}_{1:n}$ *(*$\overline{x}_{1:n}$ *are the first* n *coordinates of* $\overline{x}$*) is an approximate solution to the original linear program in the following sense*

$$c^\top \hat{x} \leq \min_{Ax=b, x\geq 0} c^\top x + LR \cdot \delta,$$

$$\|A\hat{x} - b\|_1 \leq 7n\delta \cdot \left(R \sum_{i,j} |A_{i,j}| + \|b\|_1 \right),$$

$$\hat{x} \geq 0.$$

There are different ways to construct the reduction. We pick the one such that the linear system in the modified linear program is similar to the original problem. For any diagonal matrices $\mathbf{W}_1, \mathbf{W}_2$ and any scalar α, we have

$$\mathbf{H} \stackrel{\text{def}}{=} \overline{\mathbf{A}} \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{W}_2 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \overline{\mathbf{A}}^\top = \begin{bmatrix} \mathbf{A}^\top \mathbf{W}_1 \mathbf{A} + \mathbf{D}\mathbf{W}_2\mathbf{D} & \mathbf{A}\mathbf{W}_1 u \\ (\mathbf{A}\mathbf{W}_1 u)^\top & u^\top \mathbf{W}_1 u + \alpha \end{bmatrix}.$$

Note that the second row/column block has size 1. By block inverse formula, $\mathbf{H}^{-1}v$ is an explicit formula involving $(\mathbf{A}^\top \mathbf{W}_1 \mathbf{A} + \mathbf{D}\mathbf{W}_2\mathbf{D})^{-1} v_{1:n}$ and $(\mathbf{A}^\top \mathbf{W}_1 \mathbf{A} + \mathbf{D}\mathbf{W}_2\mathbf{D})^{-1}\mathbf{A}\mathbf{W}_1 u$. Since the term $\mathbf{D}\mathbf{W}_2\mathbf{D}$ is a positive diagonal matrix, the new linear system is only shifted by a positive diagonal matrix.

With this reduction, we are ready to state our algorithm for approximate LP.

Algorithm 2: `L2InteriorPointApproximate`$(\mathbf{A}, b, c, \epsilon)$

Define the linear program $\min_{\overline{\mathbf{A}}\overline{x}=\overline{b}, \overline{x}\geq 0} \overline{c}^\top \overline{x}$ according to Lemma 8 with $\delta = \epsilon/7n$ where n is the number of columns in $\mathbf{A}$.

$(\overline{x}, \overline{s}) =$ `L2InteriorPoint`$(\overline{\mathbf{A}}, 1_{n+d+1}, 1_{n+d+1}, 1, (\epsilon/7n)^2)$.

Return $\overline{x}_{1:n}$.

Lemma 9. *Consider a linear program* $\min_{\mathbf{A}x=b, x\geq 0} c^\top x$ *with* n *variables and* d *constraints. Assume that*

1. *Diameter: For any* $x \geq 0$ *with* $\mathbf{A}x = b$*, we have that* $\|x\|_\infty \leq R$*.*

2. *Lipschitz constant of the objective:* $\|c\|_\infty \leq L$*.*

Then, `L2InteriorPointApproximate` *outputs* x *such that*

$$c^\top x \leq \min_{Ax=b, x\geq 0} c^\top x + LR \cdot \epsilon,$$

$$\|Ax - b\|_1 \leq \epsilon \cdot \left(R \sum_{i,j} |A_{i,j}| + \|b\|_1 \right),$$

$$x \geq 0.$$

The algorithm takes $O(\sqrt{n}\log(n/\epsilon))$ *steps and each step involves solving a linear system that takes* $nd^{\omega-1}$ *time. The total runtime is* $O(n^{3/2}d^{\omega-1+o(1)}\log(n/\epsilon))$*.*

Proof. By Lemma 8 with $\delta = \epsilon/7n$, it suffices to solve the modified linear program with duality gap at most δ^2. Theorem 7 shows that `L2InteriorPoint` indeed returns a pair with duality gap at most δ^2. Theorem 7 requires an initial point $xs \approx 1$ and this is satisfied according to Lemma 8.

Since `L2InteriorPoint` decreases t by $1 - \Omega(n^{-1/2})$ factor each step and since we start with $t = 1$ and ends with $t \approx \delta^2$, the number of iterations is $O(\sqrt{n}\log(n/\epsilon))$. Each step can be implemented in $nd^{\omega-1+o(1)}$. □

1.5 Rounding to Exact Solution

In this section, we discuss how to get the exact solution assuming the matrix $\mathbf{A}$, the vectors b, c are integral with bounded values.

A Deferred Proofs

Proof of Lemma 8. **Part 1.** For the first result, straightforward calculations show that $(\overline{x}, \overline{y}, \overline{s}) \in \mathbb{R}^{(n+d+1)\times(d+1)\times(n+d+1)}$ are feasible, i.e.,

$$\overline{\mathbf{A}}\overline{x} = \begin{bmatrix} \mathbf{A} & \mathbf{D} & 0_d \\ u^\top & 0_d^\top & 1 \end{bmatrix} \cdot \begin{bmatrix} 1_n \\ 1_d \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{R}b \\ U+1 \end{bmatrix} = \overline{b}$$

and

$$\overline{\mathbf{A}}^\top \overline{y} + \overline{s} = \begin{bmatrix} \mathbf{A}^\top & u \\ \mathbf{D} & 0_d \\ 0_d^\top & 1 \end{bmatrix} \cdot \begin{bmatrix} 0_d \\ -1 \end{bmatrix} + \begin{bmatrix} 1_n \\ 1_d \\ 1 \end{bmatrix} = \begin{bmatrix} -u \\ 0_d \\ -1 \end{bmatrix} + \begin{bmatrix} 1_n \\ 1_d \\ 1 \end{bmatrix} = \overline{c}.$$

Part 2. For the second result, we let

$$\mathrm{OPT} = \min_{\mathbf{A}x=b, x\geq 0} c^\top x, \quad \text{and,} \quad \overline{\mathrm{OPT}} = \min_{\overline{\mathbf{A}}\overline{x}=\overline{b}, \overline{x}\geq 0} \overline{c}^\top \overline{x}$$

For any optimal $x \in \mathbb{R}^n$ in the original LP, we consider the following $\overline{x} \in \mathbb{R}^{n+d+1}$

$$\overline{x} = \begin{bmatrix} \frac{1}{R}x \\ 0_d \\ \sum_{i\in[n]} u_i(1-\frac{x_i}{R}) + 1 \end{bmatrix}. \tag{A.1}$$

We want to argue that $\overline{x} \in \mathbb{R}^{n+d+1}$ is feasible in the modified LP. It is obvious that $\overline{x} \geq 0$, it remains to show $\overline{\mathbf{A}}\overline{x} = \overline{b} \in \mathbb{R}^{d+1}$. We have

$$\overline{\mathbf{A}}\overline{x} = \begin{bmatrix} \mathbf{A} & \mathbf{D} & 0_d \\ u^\top & 0_d^\top & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{R}x \\ 0_d \\ \sum_{i\in[n]} u_i(1-\frac{x_i}{R}) + 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{R}\mathbf{A}x \\ U+1 \end{bmatrix} = \begin{bmatrix} \frac{1}{R}b \\ U+1 \end{bmatrix} = \overline{b},$$

where the third step follows from $\mathbf{A}x = b$, and the last step follows from definition of $\overline{b}$.

Therefore, using the definition of $\overline{x}$ in (A.1), we have that

$$\overline{\mathrm{OPT}} \leq \overline{c}^\top \overline{x} = \begin{bmatrix} \frac{\delta}{L} \cdot c^\top & 1_d & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{R}x \\ 0_d \\ \sum_{i\in[n]} u_i(1-\frac{x_i}{R}) + 1 \end{bmatrix} = \frac{\delta}{LR} \cdot c^\top x = \frac{\delta}{LR} \cdot \mathrm{OPT}. \tag{A.2}$$

where the first step follows from modified program is solving a minimization problem, the second step follows from definition of $\overline{x} \in \mathbb{R}^{n+d+1}$ and $\overline{c} \in \mathbb{R}^{n+d+1}$, the last step follows from $x \in \mathbb{R}^n$ is an optimal solution in the original linear program.

Given a feasible $(\overline{x}, \overline{y}, \overline{s}) \in \mathbb{R}^{(n+d+1)\times(d+1)\times(n+d+1)}$ with duality gap δ^2, we can write $\overline{x} = \begin{bmatrix} \overline{x}_{1:n} \\ \tau \\ \theta \end{bmatrix} \in \mathbb{R}^{n+d+1}$ for some $\tau \geq 0$, $\theta \geq 0$. We can compute $\overline{c}^\top \overline{x}$ which is $\frac{\delta}{L} \cdot c^\top \overline{x}_{1:n} + \|\tau\|_1$. Then, we have

$$\frac{\delta}{L} \cdot c^\top \overline{x}_{1:n} + \|\tau\|_1 \leq \overline{\mathrm{OPT}} + \delta^2 \leq \frac{\delta}{LR} \cdot \overline{\mathrm{OPT}} + \delta^2, \tag{A.3}$$

where the first step follows from definition of duality gap, the last step follows from (A.2).

Hence, we can upper bound the $\overline{\mathrm{OPT}}$ of the transformed program as follows:

$$c^\top \hat{x} = R \cdot c^\top \overline{x}_{1:n} = \frac{LR}{\delta} \cdot \frac{\delta}{L} c^\top \overline{x}_{1:n} \leq \frac{RL}{\delta}\left(\frac{\delta}{LR} \cdot \mathrm{OPT} + \delta^2\right) = \mathrm{OPT} + LR \cdot \delta,$$

where the first step follows by $\hat{x} = R \cdot \overline{x}_{1:n}$, the third step follows by (A.3).

Note that $u^\top \overline{x}_{1:n} + \theta = U + 1$. Since $\theta \geq 0$, $\overline{x}_{1:n} \geq 0$ and $\frac{3}{2} \geq u \geq \frac{1}{2}$, we have $\|\overline{x}_{1:n}\|_1 \leq 2(U+1) \leq 5n$ and hence

$$\frac{\delta}{L} c^\top \overline{x}_{1:n} \geq -\frac{\delta}{L} \|c\|_\infty \|\overline{x}_{1:n}\|_1 \geq -\frac{\delta}{L} \|c\|_\infty 5n \geq -5\delta n, \tag{A.4}$$

where the last step follows from $\|c\|_\infty \leq L$.

We can upper bound the $\|\tau\|_1$ in the following sense,

$$\|\tau\|_1 \leq \frac{\delta}{LR} \cdot \overline{\mathrm{OPT}} + \delta^2 + 5\delta n \leq 6n\delta + \delta^2 \leq 7n\delta \tag{A.5}$$

where the first step follows from (A.3) and (A.4), the second step follows by $\mathrm{OPT} = \min_{Ax=b, x\geq 0} c^\top x \leq nLR$ (because $\|c\|_\infty \leq L$ and $\|x\|_\infty \leq R$), and the last step follows from $\delta \leq 1/2 \leq n$.

The constraint in the new polytope shows that

$$\mathbf{A}\overline{x}_{1:n} + \mathbf{D}\tau = \frac{1}{R}b.$$

Using $\hat{x} = Rx_{1:n} \in \mathbb{R}^n$, we have $\mathbf{A}\frac{1}{R}\hat{x} + \mathbf{D}\tau = \frac{1}{R}b$. Rewriting it, we have $\mathbf{A}\hat{x} - b = R\mathbf{D}\tau \in \mathbb{R}^d$ and hence

$$\|\mathbf{A}\hat{x} - b\|_1 = R\|\mathbf{D}\tau\|_1 \leq R\|\frac{1}{R}b - \mathbf{A}1_n\|_\infty \|\tau\|_1 \leq 7n\delta \cdot (R\sum_{ij} |\mathbf{A}_{ij}| + \|b\|_1),$$

where the last step follows from triangle inequality, $\|A1_n\|_\infty \leq \sum_{ij} |\mathbf{A}_{ij}|$ and (A.5). Thus, we complete the proof. $\square$