1. Idea: Rewrite the ℓ-1 norm optimization in standard LP form, find its dual problem, show that the dual can be written as problem (53), which is dual of problem (54). ~~Then~~ In other words, we show two primal problems have the same dual problem, and the optimal solution for one primal is also an optimal solution for the other primal.

1). Reformulat: $\min\ c^Tx + \mu \sum_{i=1}^{m} |a_i^Tx - b_i|$ Let $t_i \geq |a_i^Tx - b_i|, \forall i = 1, \cdots, m$

s.t. $x \geq 0$

Then we can rewrite the problem as $\min_{y \in \mathbb{R}^{n+m}} \tilde{c}^T y$ s.t. $\tilde{A}y \leq \tilde{b}$, where $y = \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+m}$,

$$\tilde{c} = \begin{bmatrix} c \\ \mu \mathbb{1} \end{bmatrix} \in \mathbb{R}^{n+m}, \quad \tilde{A} = \begin{bmatrix} -I & 0 \\ A & -I \\ -A & -I \end{bmatrix} \in \mathbb{R}^{(n+2m)\times(n+m)}, \quad \tilde{b} = \begin{bmatrix} 0 \\ b \\ -b \end{bmatrix} \in \mathbb{R}^{n+2m}$$

2). find the dual: The dual problem is $\max\ -\tilde{b}^T \begin{bmatrix} u \\ v \\ w \end{bmatrix}$ s.t. $\tilde{A}^T \begin{bmatrix} u \\ v \\ w \end{bmatrix} + \tilde{c} = 0$, $u \geq 0, v \geq 0, w \geq 0$, which can be expanded to

$$\begin{aligned} \max\ & b^T(w - v) \\ \text{s.t. } & A^T(w-v) \leq c \\ & w + v = \mu \mathbb{1} \\ & w \geq 0,\ v \geq 0 \end{aligned}$$

3). equivalence to problem (53):

- $\Rightarrow$ direction: Given feasible (w, v), simply construct $z = w - v$ and we can know that

$A^Tz = A^T(w-v) \leq c$

$\|z\|_\infty = \max_{i=1,\cdots,m} |w_i - v_i| < \max_{i=1,\cdots,m} w_i + v_i = \mu$

so we know z is a feasible point for problem (53) with same objective value.

- $\Leftarrow$ direction: Given a feasible z, construct $w \geq 0, v \geq 0$ by solving $\begin{cases} w - v = z \\ w + v = \mu \mathbb{1} \end{cases}$. Then (w, v) is also feasible with the same objective value.

2. (a). First, rewrite statement (60) to a standard form so that we can apply farkas' lemma on. Note that $w \in R^{N-1}$ but $w_i \geq 0$ for only $i = 1, \dots, M-1$, so, we introduce $w_i = u_i - v_i$, $i = M, \dots, N-1$ where $v_i \geq 0$, $u_i \geq 0$. Then we can rewrite statement (60) as

$$\exists x = \begin{bmatrix} w_1 \\ \vdots \\ w_{M-1} \\ u_M \\ \vdots \\ u_{N-1} \\ v_M \\ \vdots \\ v_{N-1} \end{bmatrix} \in R^{2N-M-1}, \quad \text{s.t. } \tilde{A} x = \tilde{b}, \; x \geq 0, \quad \text{where } \tilde{A} = \begin{bmatrix} a_2 & \cdots & a_N & -a_{M+1} & \cdots & -a_N \end{bmatrix} \in R^{n \times (2N-M-1)},$$

$$\tilde{b} = -a_1 \in R^n$$

Then we know from Farkas' lemma that its alternative is $\exists y \in R^n$ s.t. $\tilde{A}^T y \geq 0$, $\tilde{b}^T y < 0$. Let $v = -y$ and we expand them to

$$\begin{bmatrix} -a_2^T- \\ \vdots \\ -a_N^T- \\ --a_{M+1}^T- \\ \vdots \\ --a_N^T- \end{bmatrix} v \leq 0, \quad a_1^T v < 0.$$

This is

$$\begin{aligned} &a_i^T v \leq 0, \ i = 2, \dots, M \\ &a_i^T v \leq 0, \ i = M+1, \dots, N \\ &a_i^T v \geq 0, \ i = M+1, \dots, N \\ &a_1^T v < 0 \end{aligned} \Rightarrow \begin{aligned} &a_1^T v < 0 \\ &a_i^T v \leq 0, \ i = 2, \dots, M \\ &a_i^T v = 0, \ i = M+1, \dots, N \end{aligned}$$

which is exactly statement (59).

(b). Let x^*, z^* be the optimal primal-dual solution described in the problem. That is,

$a_i^T x^* = b_i$, $z_i^* = 0$, $i = 1, \dots, M$

$a_i^T x^* = b_i$, $z_i^* > 0$, $i = M+1, \dots, N$

$a_i^T x^* < b_i$, $z_i^* = 0$, $i = N+1, \dots, m$

Consider the point $\hat{x} = x^* + tv$ where $t > 0$. I claim $\hat{x}$ is also an optimal primal solution because $(\hat{x}, z^*)$ satisfies the complementary slackness condition:

For $i = 1, \dots, M$, $(b_i - a_i^T \hat{x}) z_i^* = 0$; For $i = M+1, \dots, N$, $(b_i - a_i^T \hat{x}) z_i^* = (b_i - a_i^T x^* - t a_i^T v) z_i^* = 0$

For $i = N+1, \dots, m$, $(b_i - a_i^T \hat{x}) z_i^* = 0$. Checked! (See next page)

2. (b) Continued: Now we can check that $\tilde{x}$ indeed satisfies all these conditions in (b):

$a_1^T\tilde{x} = (a_1^T x^* + t\cdot a_1^T v) = b_1 + t a_1^T v < b_1$

$a_i^T\tilde{x} = (a_i^T x^* + t\cdot a_i^T v) = b_i + t a_i^T v \le b_i$, for $i = 2, \dots, M$

$a_i^T\tilde{x} = a_i^T x^* + t\cdot a_i^T v = b_i + t a_i^T v = b_i$, for $i = M+1, \dots, N$

$a_i^T\tilde{x} = a_i^T x^* + t\cdot a_i^T v < b_i$ for sufficiently small $t > 0$, for $i = N+1, \dots, m$

(c). Similarly, construct $\tilde{z} = \begin{bmatrix} z_1^* + 1 \\ z_2^* + t\cdot w_1 \\ \vdots \\ z_M^* + t\cdot w_{M-1} \\ z_{M+1}^* + t\cdot w_M \\ \vdots \\ z_N^* + t\cdot w_{N-1} \\ z_{N+1}^* \\ \vdots \\ z_m^* \end{bmatrix}$. That is, add 1 to z_1^* to get $\tilde{z}_1$; add w to the next $N-1$ components of z^* to get $\tilde{z}_2, \dots, \tilde{z}_N$. Copy the final part to $\tilde{z}$ to get $\tilde{z}_{N+1}, \dots, \tilde{z}_m$.

It's also easy to show that $(x^*, \tilde{z})$ satisfies the complementary slackness so $\tilde{z}$ is also a dual optimal solution:

For $i = 1, \dots, M$, $(b_i - a_i^T x^*)\tilde{z}_i = 0$; For $i = M+1, \dots, N$, $(b_i - a_i^T x^*)\hat{z}_i = 0$;

For $i = N+1, \dots, m$, $(b_i - a_i^T x^*)\tilde{z}_i = (b_i - a_i^T x^*) z_i^* = 0$ checked!

Now we can show that $\tilde{z}$ satisfies all these conditions in (c):

$\tilde{z}_1 = z_1^* + 1 = 1 > 0$

$\hat{z}_i = z_i^* + tW_i = tW_i \ge 0$, for $i = 2, \dots, M$

$\hat{z}_i = z_i^* + tw_i > 0$ for sufficient small t, for $i = M+1, \dots, N$

$\hat{z}_i = z_i^* = 0$, for $i = N+1, \dots, m$

(d). From part (a) we know that exactly one of the two alternatives is true. If alternative (I) is true, then we can come up with a new optimal primal-dual pair $(\hat{x}, z^*)$ with at most $\hat{M} = M-1 < M$ common zeros because $a_1^T \hat{x} < b_1$. Similarly, if alternative (II) is true, then we can construct a new optimal primal-dual pair $(x^*, \hat{z})$ with at most $\hat{M} = M-1 < M$ common zeros because $\hat{z}_1 > 0$.
Then, if $\hat{M} = 0$, we're done. If $\hat{M} > 0$, we just need to apply a similar argument to construct a new optimal primal-dual pair with less common zeros. We keep iterating this process until we find an optimal primal-dual pair with zero common zeros. Done.

3. Extreme points $\hat{x}$ of P are unique solutions of n linearly independent tight inequalities. Let A' = active constriants $\in \mathbb{R}^{n \times n}$ and b' is the corresponding values for those constraints. I.e. $A'x = b'$. Cramer's Rule told us that the solution of this equation can be written as $x_i = \dfrac{\det(A_i')}{\det(A')}$ for $i = 1, \dots, n$ where A_i' is A' with ~~col~~ i^{th} column replaced by b'.
The denominator $\det(A')$ is either 1 or -1 because ~~A is un~~ A is totally unimodular, and A_i' has only integer entries, so $\det(A_i')$ is also an interger.
Hence, x will be an integer vector.

4. Mathematical Induction:
1) Base case: when $k=1$, the determinant of any 1×1 matrix is either 1, 0, or -1 because an incidence matrix has only entries of 1, 0, or -1.

(see next page)

2). Induction step: Suppose we know all sub-matrices of size $k \times k$ have determinant 1, 0, or -1. Then we want to show this is also true for all $(k+1)\times(k+1)$ matrices. Say A_{sub} is a $(k+1)\times(k+1)$ matrix. If there is

$$A_{sub} = \underbrace{\left.\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\right\} k+1}_{k+1}$$

a column in A_{sub} with all zeros then the determinant is zero; If there is a column in A_{sub} with only one 1 (or only one -1), then we can compute the determinant of A_{sub} using cofactor formula along that column, which by induction is either 1, 0, or -1.

If all columns in A_{sub} have exactly one 1 and one -1, then we know that in this case the sum of all rows is zero because each component is ~~e~~just $1-1=0$. This means those $k+1$ rows are linearly dependent and hence the determinant is zero.

Hence, we have proved that all $(k+1)\times(k+1)$ sub-matrices have determinant of 1, 0, or -1.

Therefore, ~~sa~~ all squre sub-matrices in an incidence matrix has determinant 1, 0, or -1.

Thus, the incidence matrix is totally unimodular.