Lecture 8. Duality of Linear Programming

8.1 Primal and dual problems

Consider the LP:

$$\begin{aligned} \min \quad & 3x_1 + 2x_2 \\ \text{subject to} \quad & x_1 + x_2 \le 5\,, \\ & x_1 \le 3\,, \\ & x_1, x_2 \ge 0\,. \end{aligned}$$

Trivially $x_1 = x_2 = 0$ is optimal. How about $\max 3x_1 + 2x_2$?
It is also easy to see that $3x_1 + 2x_2 \le 13$ since

$$3x_1 + 2x_2 = 2(x_1 + x_2) + x_1 \le 5 \times 2 + 3.$$

In general, given $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and the LP constraints $Ax \le b, x \ge 0$, we can assign each constraint a coefficient y_i as follows:

$$\begin{array}{rlr} \max & c_1x_1 + c_2x_2 + \cdots + c_nx_n & \\ \text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1\,. & y_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2\,. & y_2 \\ & \vdots & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m\,. & y_m \end{array}$$

If $y_j \ge 0$ and $\sum_i y_i a_{ij} \ge c_j$ for all j, then $\sum_i y_i b_i$ is an upper bound, which further implies that

$$\max_{Ax \le b, x \ge 0} c_1x_1 + c_2x_2 + \cdots + c_nx_n \le \min_{A^\top y \ge c, y \ge 0} y_1b_1 + y_2b_2 + \cdots + y_mb_m\,.$$

We want the upper bound as small as possible. Therefore, we would like to solve the following LP:

$$\begin{aligned} \min \quad & \boldsymbol{b}^\top \boldsymbol{y} \\ \text{subject to} \quad & \boldsymbol{A}^\top \boldsymbol{y} \ge \boldsymbol{c}\,, \\ & \boldsymbol{y} \ge \boldsymbol{0}\,, \end{aligned}$$

which is termed as the *dual linear program*.

Proposition

The dual of the dual is the primal.

Proof

The dual program can be rewritten as:

$$\begin{aligned} \max \quad & -\boldsymbol{b}^\top \boldsymbol{y} \\ \text{subject to} \quad & -\boldsymbol{A}^\top \boldsymbol{y} \le -\boldsymbol{c}\,, \\ & \boldsymbol{y} \ge \boldsymbol{0}\,. \end{aligned}$$

It is clear that the dual of dual is:

$$\begin{aligned} \min \quad & -\boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & -\boldsymbol{A}^\top \boldsymbol{x} \ge -\boldsymbol{b}\,, \\ & \boldsymbol{x} \ge \boldsymbol{0}\,, \end{aligned}$$

which is equivalent to the primal.

We note that the primal problem may have different forms. If the primal has a constraint $a_i^\top \boldsymbol{x} \ge b_i$, then we have the corresponding variable $y_i \le 0$ in the dual. If the primal has a constraint $\boldsymbol{a}_i^\top \boldsymbol{x} = b_i$, then we have the corresponding variable y_i unconstrained in sign in the dual.

Now we use the form $\max\ \boldsymbol{c}^\top \boldsymbol{x}$, subject to $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$, $\boldsymbol{x} \ge \boldsymbol{0}$. By the discussion above, we know that the optimal value of the dual problem gives an upper bound of the primal problem. This property is called the *weak duality*.

Theorem (*Weak duality theorem*)

If $\boldsymbol{x}$ is feasible for primal and $\boldsymbol{y}$ is feasible for dual, then $\boldsymbol{c}^\top \boldsymbol{x} \le \boldsymbol{b}^\top \boldsymbol{y}$.

Proof

Since $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$, $\boldsymbol{A}^\top \boldsymbol{y} \ge \boldsymbol{c}$ and $\boldsymbol{x}, \boldsymbol{y} \ge \boldsymbol{0}$, we have $\boldsymbol{c}^\top \boldsymbol{x} \le \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{y}^\top \boldsymbol{b}$.

Corollary

If $\boldsymbol{c}^\top \boldsymbol{x}$ is unbounded above, then the dual problem is infeasible.

However, if the primal problem is infeasible, we cannot conclude that the dual problem is unbounded below. It is possible that both primal and dual problems are infeasible.

Example

Consider the following problem:

$$\begin{aligned} \max \quad & 2x_1 - x_2 \\ \text{subject to} \quad & x_1 - x_2 \leq 1\,, \\ & x_2 - x_1 \leq -2\,, \\ & x_1, x_2 \geq 0\,. \end{aligned}$$

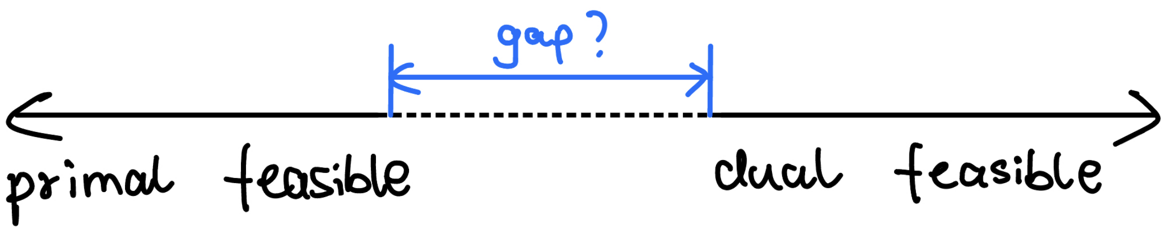
Its dual problem is

$$\begin{aligned} \min \quad & y_1 - 2y_2 \\ \text{subject to} \quad & y_1 - y_2 \leq 1\,, \\ & y_2 - y_1 \leq -2\,, \\ & y_1, y_2 \geq 0\,. \end{aligned}$$

It is easy to check that both of them are infeasible.

8.2 Strong duality

We've already known that any feasible solution of the dual gives an upper to the primal. The question is that, is there any gap between the optimal value of the primal and the optimal value of the dual?



Theorem (*Strong duality theorem*)

If the primal problem has a finite optimal solution $\boldsymbol{x}^*$, then the dual problem also has a finite optimal solution $\boldsymbol{y}^*$ with the same optimal value of the primal.

Namely, it always holds that $\boldsymbol{c}^\top \boldsymbol{x}^* = \boldsymbol{b}^\top \boldsymbol{y}^*$.

Now we can complete the following table.

	unbounded	infeasible	$\exists$ optimal
unbounded	×	√	×
infeasible	√	√	×
$\exists$ optimal	×	×	√

The proof of the strong duality is an application of the *Farkas' lemma*, which we have introduced in Lecture 4.

Theorem (*Farkas' lemma*)

Let $\boldsymbol{A} \in \mathbb{R}^{n\times m}, \boldsymbol{b} \in \mathbb{R}^n$. Then exactly one of the following sets is empty:

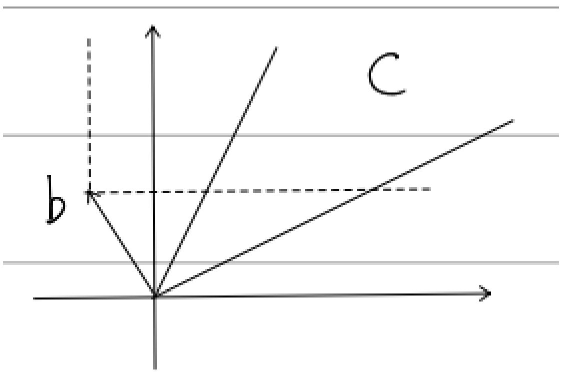
1. $\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$;
2. $\{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{A}^\top \boldsymbol{y} \leq \boldsymbol{0}, \boldsymbol{b}^\top \boldsymbol{y} > 0\}$.

To apply this lemma, we consider the following corollary.

Corollary

Suppose $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then exactly one of the followings is true:

1. There exists $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$.
2. There exists $\boldsymbol{y} \in \mathbb{R}^m$ such that $\boldsymbol{A}^\top \boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{b}^\top \boldsymbol{y} < 0$, and $\boldsymbol{y} \leq \boldsymbol{0}$.



Let C be the cone of column vectors of $\boldsymbol{A}$. The corollary tells us that either C

intersects the region $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \geq \boldsymbol{b}\}$, or there exists a hyperplane $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{y}^\mathsf{T}\boldsymbol{x} = 0\}$ strictly separating $\boldsymbol{b}$ and C, where $\boldsymbol{b}^\mathsf{T}\boldsymbol{y} < 0$ and $\boldsymbol{y} \leq \boldsymbol{0}$.

Proof of the corollary

Let $\boldsymbol{A}' = (\boldsymbol{A} \quad -\boldsymbol{I}) \in \mathbb{R}^{m\times(m+n)}$. Applying the Farkas' lemma to $\boldsymbol{A}'$ and $\boldsymbol{b}$, it gives that exactly one of the followings is true:

1. There exists $\boldsymbol{x}' \in \mathbb{R}^{m+n}_{\geq 0}$ such that $\boldsymbol{A}'\boldsymbol{x}' = b$;
2. There exists $\boldsymbol{y} \in \mathbb{R}^m$ such that $\boldsymbol{A}'^\mathsf{T}\boldsymbol{y} \geq \boldsymbol{0}$ and $\boldsymbol{b}^\mathsf{T}\boldsymbol{y} < 0$.

Note that item 1 is equivalent to $\exists\, \boldsymbol{x} \in \mathbb{R}^n_{\geq 0}$ such that $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$, and item 2 is equivalent to $\boldsymbol{A}^\mathsf{T}\boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{y} \leq \boldsymbol{0}$, and $\boldsymbol{b}^\mathsf{T}\boldsymbol{y} < 0$. So we are done.

We are now ready to prove the strong duality theorem.

Proof of the strong duality theorem

Without loss of generality, we assume that the dual problem has an optimal solution $\boldsymbol{y}^*$ (instead of the primal problem in the statement). Suppose the strong duality is not true. Then there does not exist feasible solution $\boldsymbol{x}$ of the primal such that $\boldsymbol{c}^\mathsf{T}\boldsymbol{x} = \boldsymbol{b}^\mathsf{T}\boldsymbol{y}^*$.
Let $\gamma = \boldsymbol{b}^\mathsf{T}\boldsymbol{y}^*$. It is equivalent to that there does not exists $\boldsymbol{x} \in \mathbb{R}^n_{\geq 0}$ such that

$$\begin{pmatrix} -\boldsymbol{A} \\ \boldsymbol{c}^\mathsf{T} \end{pmatrix} \boldsymbol{x} \geq \begin{pmatrix} -\boldsymbol{b} \\ \gamma \end{pmatrix}.$$

Then the corollary of the Farkas' Lemma shows that there exists $\boldsymbol{y} \in \mathbb{R}^m_{\leq 0}$ and $w \in \mathbb{R}_{\leq 0}$ such that

$$(-\boldsymbol{A}^\mathsf{T} \quad \boldsymbol{c}) \begin{pmatrix} \boldsymbol{y} \\ w \end{pmatrix} \geq \boldsymbol{0}\,, \quad \text{and} \quad (-\boldsymbol{b}^\mathsf{T} \quad \gamma) \begin{pmatrix} \boldsymbol{y} \\ w \end{pmatrix} < \boldsymbol{0}\,.$$

Now we can claim that $\boldsymbol{y}^*$ is not an optimal solution to the dual problem, which contradicts to our assumption. Consider the following two cases.

- Case 1. $w = 0$. Then we have $(-\boldsymbol{A}^\mathsf{T} \quad \boldsymbol{c}) \begin{pmatrix} \boldsymbol{y} \\ w \end{pmatrix} = -\boldsymbol{A}^\mathsf{T}\boldsymbol{y} + w\boldsymbol{c} = -\boldsymbol{A}^\mathsf{T}\boldsymbol{y}$ and $(-\boldsymbol{b}^\mathsf{T} \quad \gamma) \begin{pmatrix} \boldsymbol{y} \\ w \end{pmatrix} = -\boldsymbol{b}^\mathsf{T}\boldsymbol{y} + \gamma w = -\boldsymbol{b}^\mathsf{T}\boldsymbol{y}$. Thus, $-\boldsymbol{A}^\mathsf{T}\boldsymbol{y} \geq \boldsymbol{0}$ and $-\boldsymbol{b}^\mathsf{T}\boldsymbol{y} < 0$.
  Noting that

  $$\boldsymbol{A}^\mathsf{T}(\boldsymbol{y}^* - \boldsymbol{y}) = \boldsymbol{A}^\mathsf{T}\boldsymbol{y} - \boldsymbol{A}^\mathsf{T}\boldsymbol{y} \geq \boldsymbol{A}^\mathsf{T}\boldsymbol{y}^* \geq \boldsymbol{c}\,, \quad \text{and} \quad \boldsymbol{y}^* - \boldsymbol{y} \geq \boldsymbol{y}^* \geq \boldsymbol{0}\,,$$

$\boldsymbol{y}^* - \boldsymbol{y}$ is also a feasible solution to the dual problem. But $\boldsymbol{b}^\top(\boldsymbol{y}^* - \boldsymbol{y}) < \boldsymbol{b}^\top \boldsymbol{y}^*$.

- Case 2. $w < 0$. Dividing w on both sides, it leads to

$$(-\boldsymbol{A}^\top \quad \boldsymbol{c}) \begin{pmatrix} \boldsymbol{y}/w \\ 1 \end{pmatrix} \le \boldsymbol{0}\,, \quad \text{and} \quad (-\boldsymbol{b}^\top \quad \gamma) \begin{pmatrix} \boldsymbol{y}/w \\ 1 \end{pmatrix} > \boldsymbol{0}\,,$$

  which implies that $\boldsymbol{A}^\top(\boldsymbol{y}/w) \ge \boldsymbol{c}$ and $\boldsymbol{b}^\top(\boldsymbol{y}/w) < \gamma$. So $\boldsymbol{y}/w$ is a feasible solution to the dual problem, and better than $\boldsymbol{y}^*$.

The proof is concluded by the contradictions in both cases.

In the proof of the corollary of the Farkas' lemma, we employ the matrix $(\boldsymbol{A} \quad -\boldsymbol{I})$, which looks similar to the standard form of the linear program. In fact, if we consider the standard form, the proof of the strong duality can apply the Farkas' lemma directly, instead of using the corollary.

Complementary slackness

As an application of the strong duality of linear program, the following theorem reveals some relations between the optimal solutions to the primal and the dual.

Theorem (*Complementary slackness*)

Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are feasible solutions to the primal problem and the dual problem, respectively. Then $\boldsymbol{x}$ and $\boldsymbol{y}$ are optimal solutions if and only if

$$\boldsymbol{y}^\top(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) = 0\,, \qquad \boldsymbol{x}^\top(\boldsymbol{A}^\top\boldsymbol{y} - \boldsymbol{c}) = 0\,.$$

The complementary slackness shows that if the i-th constraint of the primal at the optimal solution $\boldsymbol{x}$ is not tight, then the corresponding variable y_i is 0 in the optimal solution of the dual, and vice versa. Namely, we have that

$$\begin{cases} \text{either } y_i = 0, \text{ or } (\boldsymbol{A}\boldsymbol{x})_i = b_i \text{ is tight,} \\ \text{either } x_j = 0, \text{ or } (\boldsymbol{A}^\top\boldsymbol{y})_j = c_j \text{ is tight.} \end{cases}$$

Proof

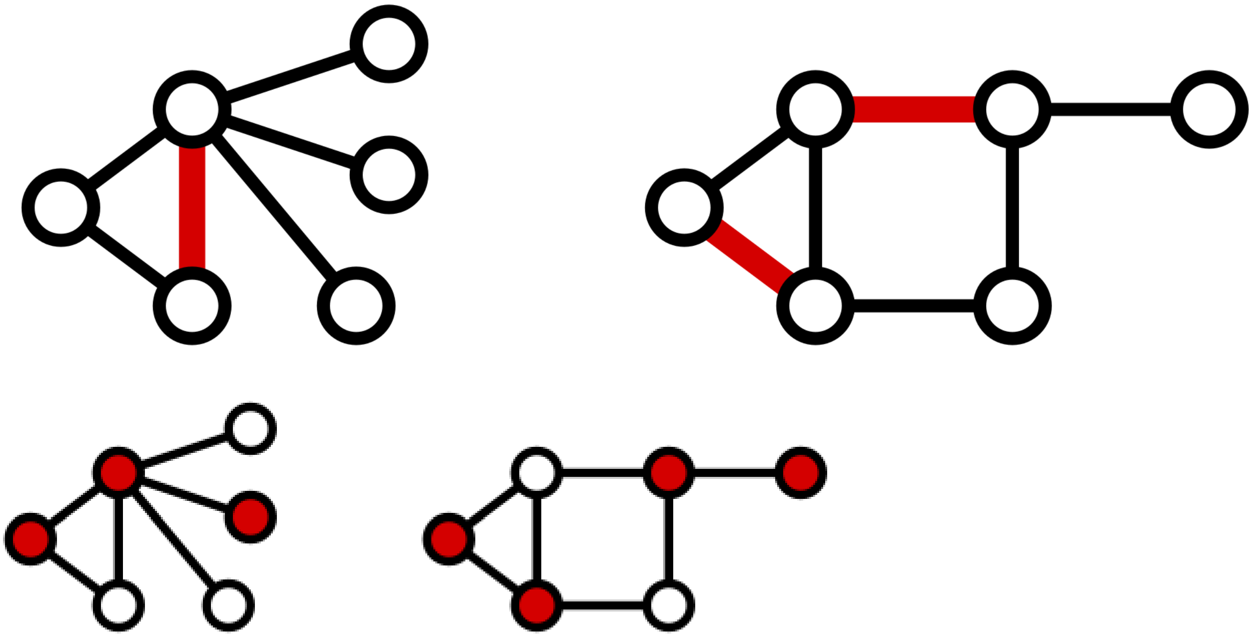
It is clear that $\boldsymbol{c}^\top\boldsymbol{x} \le \boldsymbol{y}^\top\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{y}^\top\boldsymbol{b}$. Then $\boldsymbol{x}$ and $\boldsymbol{y}$ are optimal solutions if and only if $\boldsymbol{c}^\top\boldsymbol{x} = \boldsymbol{y}^\top\boldsymbol{b}$, that is,

$$\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x}\,, \quad \text{and} \quad \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}\,.$$

8.3 Applications of linear programming duality

We now consider an application of the strong duality in linear programming.

Given an undirected graph $G = (V, E)$, a *matching* is a set of edges such that no edge in the set is a loop and no two edges share common vertices, and a *vertex cover* is a set of vertices that includes at least one endpoint of every edge. Here are some examples of matchings and vertex covers.



Clearly the set $\emptyset$ is a matching and V is a vertex cover. So we consider the problem of *maximum matching* and *minimum vertex cover*.

Theorem (Kőnig's theorem)

Suppose the graph $G = (V, E)$ is a *bipartite graph*, namely, $V = L \cup R$ for some disjoint L and R, such that $E \subseteq L \times R$. Then the size of its maximum matching equals the size of its minimum vertex cover.

We can write an integer programming formulation of the maximum matching problem. Any matching $M \subseteq E$ can be represented by $|E|$ variables such that $x_e = 1$ if the edge $e \in M$ and $x_e = 0$ otherwise. Conversely, any assignment to $\{x_e\}_{e \in E}$ can represent a matching if $x_e \in \{0, 1\}$ and $\sum_{(u,v) \in E} x_{(u,v)} \leq 1$ for all $v \in V$.

So the problem can be formulated as

$$\begin{aligned} \max \quad & \sum_{e\in E} x_e \\ \text{subject to} \quad & \forall\, e \in E,\ x_e \in \{0,1\}\,; \\ & \forall\, v \in V,\ \sum_{(u,v)\in E} x_{(u,v)} \le 1\,. \end{aligned}$$

If we relax the constraints $x_e \in \{0,1\}$ to be $x_e \ge 0$, it can be reformulated as a linear program

$$\begin{aligned} \max \quad & \sum_{e\in E} x_e \\ \text{subject to} \quad & \forall\, e \in E,\ x_e \ge 0\,; \\ & \forall\, v \in V,\ \sum_{(u,v)\in E} x_{(u,v)} \le 1\,. \end{aligned}$$

The relaxed problem is called the *maximum fractional matching problem*.

Analogously, for each vertex v we can assign a variable y_v to represent whether v is in the vertex cover. We relax the constraints $y_v \in \{0,1\}$ again. Then we obtain the problem of *maximum fractional vertex cover* as follows

$$\begin{aligned} \min \quad & \sum_{v\in V} y_v \\ \text{subject to} \quad & \forall\, v \in V,\ y_v \ge 0\,; \\ & \forall\, (u,v) \in E,\ y_u + y_v \ge 1\,. \end{aligned}$$

It is easy to verify that the fractional minimum vertex cover problem is the dual of the fractional maximum matching problem. So we know that for any graph, the size of the maximum fractional matching equals the size of the minimum fractional vertex cover.

In a bipartite graph G, we can further show that the size of the maximum fractional matching equals the size of the maximum matching. Given a fractional matching $\{x_e\}_{e\in E}$, consider the subgraph consisting of fractional edges e where $x_e \notin \{0,1\}$.

- Case 1. There exists a cycle $\{v_1, v_2, \ldots, v_\ell\}$. Note that ℓ is an even number since G is bipartite. Let $\varepsilon = \min\{1 - x_{(v_1,v_2)}, x_{(v_2,v_3)}, 1 - x_{(v_3,v_4)}, x_{(v_4,v_5)}, \ldots, 1 - x_{(v_{\ell-1},v_\ell)}, x_{(v_\ell,v_1)}\}$. Then add ε to $x_{(v_i,v_{i+1})}$ for all odd i, and subtract ε from $x_{(v_i,v_{i+1})}$ for all even i (we assume that $v_{\ell+1} = v_1$). The resulting $\{x_e\}$ satisfy all constraints and the size of the fractional matching remains the same.

- Case 2. There is no cycles. Then choose any path $\{v_1, \ldots, v_\ell\}$. Note that all edges e incident to v_1 has $x_e \in \{0,1\}$ except $x_{(v_1,v_2)}$. So $x_e = 0$ if v_1 belongs to e but $e \neq (v_1, v_2)$. Similarly, $x_e = 0$ if v_ℓ belongs to e but $e \neq (v_{\ell-1}, v_\ell)$. Again, let $\varepsilon = \min\{1 - x_{(v_1,v_2)}, x_{(v_2,v_3)}, 1 - x_{(v_3,v_4)}, x_{(v_4,v_5)}, \ldots, x_{(v_{\ell-1},v_\ell)}$ or $1 - x_{(v_{\ell-1},v_\ell)}\}$. Then subtract ε from $x_{(v_i,v_{i+1})}$ for all odd i, and add ε to $x_{(v_i,v_{i+1})}$ for all even i (we assume that $v_{\ell+1} = v_1$). Now the resulting $\{x_e\}$ satisfy all constraints and the size of the fractional matching is nondecreasing.
  Each operation decrease the number of fractional edges. So there is no fractional edges after finite many operations. Consequently, any fractional matching can be converted into an integral matching that is not worse, which implies that the size of the maximum fractional matching equals the size of the maximum matching

We can also show that the size of the minimum fractional vertex cover equals the size of the minimum vertex cover in any bipartite graph G. Suppose $G = (L \cup R, E)$ where $L \cap R = \emptyset$ and $E \subseteq L \times R$. We construct a random vertex cover C as follows.
Uniformly choose a real number $p \in [0,1]$. For every $u \in L$, let $u \in C$ if $0 \le p \le y_u$, and $u \notin C$ otherwise. For every $v \in L$, let $v \in C$ if $1 - y_v \le p \le 1$, and $v \notin C$ otherwise. For every $u \in L$ and $v \in R$, if $(u, v) \in E$, then $y_u + y_v \ge 1$. So at least one of $\{u, v\}$ is in C, which gives that C is a vertex cover.
Now we calculate the *expected size* of C. For any $v \in L \cup R$, $\Pr[v \in C] = y_v$. Thus, by the linearity of expectation, $\mathbb{E}[|C|] = \sum_{v \in L \cup R} y_v$, which is the size of the minimum fractional vertex cover. In addition, there exists $p \in [0,1]$ such that the vertex cover C' constructed by p has the size $|C'| \le \mathbb{E}[|C|]$.

Overall, we conclude that in any bipartite graph, the size of the maximum matching equals the size of the minimum vertex cover.