For any optimal solution y^* of (D), since v is a feasible solution of (D) such that $vb < \mu + \epsilon$, we must have

$$\mu \leq y^* b < \mu + \epsilon,$$

and since our reasoning is valid for *any* $\epsilon > 0$, we conclude that $cx^* = \mu = y^* b$.

If we assume that the dual program (D) has a feasible solution and is bounded below, since the dual of (D) is (P), we conclude that (P) is also feasible and bounded above. $\square$

The strong duality theorem can also be proven by the simplex method, because when it terminates with an optimal solution of (P), the final tableau also produces an optimal solution y of (D) that can be read off the reduced costs of columns $n+1, \ldots, n+m$ by flipping their signs. We follow the proof in Ciarlet [41] (Chapter 10).

Theorem 47.8. *Consider the Linear Program (P),*

$$\begin{aligned} &\text{maximize} \quad cx \\ &\text{subject to} \quad Ax \leq b \ \textit{and}\ x \geq 0, \end{aligned}$$

its equivalent version $(P2)$ in standard form,

$$\begin{aligned} &\text{maximize} \quad \widehat{c}\,\widehat{x} \\ &\text{subject to} \quad \widehat{A}\widehat{x} = b \ \textit{and}\ \widehat{x} \geq 0, \end{aligned}$$

where $\widehat{A}$ is an $m \times (n+m)$ matrix, $\widehat{c}$ is a linear form in $(\mathbb{R}^{n+m})^*$, and $\widehat{x} \in \mathbb{R}^{n+m}$, given by

$$\widehat{A} = \begin{pmatrix} A & I_m \end{pmatrix}, \quad \widehat{c} = \begin{pmatrix} c & 0_m^\top \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \overline{x} = \begin{pmatrix} x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix}, \quad \widehat{x} = \begin{pmatrix} x \\ \overline{x} \end{pmatrix},$$

and the Dual (D) of (P) given by

$$\begin{aligned} &\text{minimize} \quad yb \\ &\text{subject to} \quad yA \geq c \ \textit{and}\ y \geq 0, \end{aligned}$$

where $y \in (\mathbb{R}^m)^*$. If the simplex algorithm applied to the Linear Program $(P2)$ terminates with an optimal solution $(\widehat{u}^*, K^*)$, where $\widehat{u}^*$ is a basic feasible solution and K^* is a basis for $\widehat{u}^*$, then $y^* = \widehat{c}_{K^*}\widehat{A}_{K^*}^{-1}$ is an optimal solution for (D) such that $\widehat{c}\,\widehat{u}^* = y^* b$. Furthermore, y^* is given in terms of the reduced costs by $y^* = -((\overline{c}_{K^*})_{n+1} \ \ldots (\overline{c}_{K^*})_{n+m})$.

Proof. We know that K^* is a subset of $\{1, \ldots, n+m\}$ consisting of m indices such that the corresponding columns of $\widehat{A}$ are linearly independent. Let $N^* = \{1, \ldots, n+m\} - K^*$. The simplex method terminates with an optimal solution in Case (A), namely when

$$\widehat{c}_j - \sum_{k \in k} \gamma_k^j \widehat{c}_k \leq 0 \quad \text{for all } j \in N^*,$$