Let us reformulate the problem as

$$\begin{aligned} &\text{minimize} \quad f(y) \\ &\text{subject to} \\ &\qquad\qquad Ax + b = y, \end{aligned}$$

where we introduced the new variable $y \in \mathbb{R}^m$ and the equality constraint $Ax + b = y$. The two problems are obviously equivalent. The Lagrangian of the reformulated problem is

$$L(x, y, \mu) = f(y) + \mu^\top (Ax + b - y)$$

where $\mu \in \mathbb{R}^m$. To find the dual function $G(\mu)$ we minimize $L(x, y, \mu)$ over x and y. Minimizing over x we see that $G(\mu) = -\infty$ unless $A^\top \mu = 0$, in which case we are left with

$$G(\mu) = b^\top \mu + \inf_y (f(y) - \mu^\top y) = b^\top \mu - \inf_y (\mu^\top y - f(y)) = b^\top \mu - f^*(\mu),$$

where f^* is the conjugate of f. It follows that the dual program can be expressed as

$$\begin{aligned} &\text{maximize} \quad b^\top \mu - f^*(\mu) \\ &\text{subject to} \\ &\qquad\qquad A^\top \mu = 0. \end{aligned}$$

This formulation of the dual is much more useful than the dual of the original program.

Example 50.12. As a concrete example, consider the following unconstrained program:

$$\text{minimize} \quad f(x) = \log\left(\sum_{i=1}^n e^{(A^i)^\top x + b_i}\right)$$

where A^i is a column vector in $\mathbb{R}^n$. We reformulate the problem by introducing new variables and equality constraints as follows:

$$\begin{aligned} &\text{minimize} \quad f(y) = \log\left(\sum_{i=1}^n e^{y_i}\right) \\ &\text{subject to} \\ &\qquad\qquad Ax + b = y, \end{aligned}$$

where A is the $n \times n$ matrix whose columns are the vectors A^i and $b = (b_1, \ldots, b_n)$. Since by Example 50.8(8), the conjugate of the log-sum-exp function $f(y) = \log\left(\sum_{i=1}^n e^{y_i}\right)$ is

$$f^*(\mu) = \begin{cases} \sum_{i=1}^n \mu_i \log \mu_i & \text{if } \mathbf{1}^\top \mu = 1 \text{ and } \mu \geq 0 \\ \infty & \text{otherwise,} \end{cases}$$