If the conditions of Theorem 50.19(1) hold, which in our case means that for every $\lambda \in \mathbb{R}^m$, there is a unique $u_\lambda \in \mathbb{R}^n$ such that

$$G(\lambda) = L(u_\lambda, \lambda) = \inf_{u \in \mathbb{R}^n} L(u, \lambda),$$

and that the function $\lambda \mapsto u_\lambda$ is continuous, then G is differentiable. Furthermore, we have

$$\nabla G_\lambda = Au_\lambda - b,$$

and for any solution $\mu = \lambda^*$ of the dual problem

$$\begin{aligned} &\text{maximize} && G(\lambda) \\ &\text{subject to} && \lambda \in \mathbb{R}^m, \end{aligned}$$

the vector $u^* = u_\mu$ is a solution of the primal Problem (P). Furthermore, $J(u^*) = G(\lambda^*)$, that is, the duality gap is zero.

The dual ascent method is essentially gradient descent applied to the dual function G. But since G is maximized, gradient descent becomes gradient ascent. Also, we no longer worry that the minimization problem $\inf_{u \in \mathbb{R}^n} L(u, \lambda)$ has a unique solution, so we denote by u^+ some minimizer of the above problem, namely

$$u^+ = \arg\min_u L(u, \lambda).$$

Given some initial dual variable λ^0, the *dual ascent method* consists of the following two steps:

$$\begin{aligned} u^{k+1} &= \arg\min_u L(u, \lambda^k) \\ \lambda^{k+1} &= \lambda^k + \alpha^k (Au^{k+1} - b), \end{aligned}$$

where $\alpha^k > 0$ is a step size. The first step is used to compute the "new gradient" (indeed, if the minimizer u^{k+1} is unique, then $\nabla G_{\lambda^k} = Au^{k+1} - b$), and the second step is a dual variable update.

Example 52.1. Let us look at a very simple example of the gradient ascent method applied to a problem we first encountered in Section 42.1, namely minimize $J(x, y) = (1/2)(x^2 + y^2)$ subject to $2x - y = 5$. The Lagrangian is

$$L(x, y, \lambda) = \frac{1}{2}(x^2 + y^2) + \lambda(2x - y - 5).$$

See Figure 52.1.

The method of Lagrangian duality says first calculate

$$G(\lambda) = \inf_{(x,y) \in \mathbb{R}^2} L(x, y, \lambda).$$