Example 42.1. For instance, we may want to minimize the quadratic function

$$Q(x_1, x_2) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$$

subject to the constraint

$$2x_1 - x_2 = 5.$$

The solution for which $Q(x_1, x_2)$ is minimum is no longer $(x_1, x_2) = (0, 0)$, but instead, $(x_1, x_2) = (2, -1)$, as will be shown later.

Geometrically, the graph of the function defined by $z = Q(x_1, x_2)$ in $\mathbb{R}^3$ is a paraboloid of revolution P with axis of revolution Oz. The constraint

$$2x_1 - x_2 = 5$$

corresponds to the vertical plane H parallel to the z-axis and containing the line of equation $2x_1 - x_2 = 5$ in the xy-plane. Thus, as illustrated by Figure 42.1, the constrained minimum of Q is located on the parabola that is the intersection of the paraboloid P with the plane H.

A nice way to solve constrained minimization problems of the above kind is to use the method of *Lagrange multipliers* discussed in Section 40.1. But first let us define precisely what kind of minimization problems we intend to solve.

Definition 42.3. The *quadratic constrained minimization problem* consists in minimizing a quadratic function

$$Q(x) = \frac{1}{2}x^\top A^{-1}x - b^\top x$$

subject to the linear constraints

$$B^\top x = f,$$

where A^{-1} is an $m \times m$ symmetric positive definite matrix, B is an $m \times n$ matrix of rank n (so that $m \geq n$), and where $b, x \in \mathbb{R}^m$ (viewed as column vectors), and $f \in \mathbb{R}^n$ (viewed as a column vector).

The reason for using A^{-1} instead of A is that the constrained minimization problem has an interpretation as a set of equilibrium equations in which the matrix that arises naturally is A (see Strang [169]). Since A and A^{-1} are both symmetric positive definite, this doesn't make any difference, but it seems preferable to stick to Strang's notation.

In Example 42.1 we have $m = 2$, $n = 1$,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad f = 5.$$

As explained in Section 40.1, the method of Lagrange multipliers consists in incorporating the n constraints $B^\top x = f$ into the quadratic function $Q(x)$, by introducing extra variables