Chapter 9

Lagrange multipliers. Portfolio optimization.

The Lagrange multipliers method for finding constrained extrema of multivariable functions.

9.1 Lagrange multipliers

Optimization problems often require finding extrema of multivariable functions subject to various constraints. One method to solve such problems is by using Lagrange multipliers, as outlined below.

Let $U \subset \mathbb{R}^n$ be an open set, and let $f : U \to \mathbb{R}$ be a smooth function, e.g., infinitely many times differentiable. We want to find the extrema of $f(x)$ subject to m constraints given by $g(x) = 0$, where $g : U \to \mathbb{R}^m$ is a smooth function, i.e.,

Find $x_0 \in U$ such that

$$\max_{\substack{g(x)=0\\ x \in U}} f(x) = f(x_0) \quad \text{or} \quad \min_{\substack{g(x)=0\\ x \in U}} f(x) = f(x_0). \tag{9.1}$$

Problem (9.1) is called a constrained optimization problem. For this problem to be well posed, a natural assumption is that the number of constraints is smaller than the number of the degrees of freedom, i.e., $m < n$.

A point $x_0 \in U$ satisfying (9.1) is called a constrained extremum point of the function $f(x)$ with respect to the constraint function $g(x)$.

To solve the constrained optimization problem (9.1), let $\lambda = (\lambda_i)_{i=1:m}$ be a column vector of the same size, m, as the number of constraints; λ is called the Lagrange multipliers vector.

The Lagrangian associated to problem (9.1) is the function $F : U \times \mathbb{R}^m \to \mathbb{R}$ given by

$$F(x, \lambda) = f(x) + \lambda^t g(x). \tag{9.2}$$

If $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}$, then $F(x,\lambda)$ can be written explicitly as

$$F(x,\lambda) \;=\; f(x) \;+\; \sum_{i=1}^{m} \lambda_i g_i(x).$$

Example: For a better understanding of the Lagrange multipliers method, in parallel with presenting the general theory, we will solve the following constrained optimization problem:

Find the maximum and minimum values of the function $f(x_1,x_2,x_3) = 4x_2 - 2x_3$ subject to the constraints $2x_1 = x_2 + x_3$ and $x_1^2 + x_2^2 = 13$.

We reformulate the problem by introducing the functions $f : \mathbb{R}^3 \to \mathbb{R}$ and $g : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$f(x) = 4x_2 - 2x_3; \quad g(x) = \begin{pmatrix} 2x_1 - x_2 - x_3 \\ x_1^2 + x_2^2 - 13 \end{pmatrix}, \tag{9.3}$$

where $x = (x_1,x_2,x_3)$. We want to find $x_0 \in \mathbb{R}^3$ such that

$$\max_{\substack{g(x)=0 \\ x\in\mathbb{R}^3}} f(x) \;=\; f(x_0) \qquad \text{or} \qquad \min_{\substack{g(x)=0 \\ x\in\mathbb{R}^3}} f(x) \;=\; f(x_0).$$

Let $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$ be the Lagrange multiplier. The Lagrangian associated to this problem is

$$\begin{aligned} F(x,\lambda) \;&=\; f(x) \;+\; \lambda_1 g_1(x) \;+\; \lambda_2 g_2(x) \\ &=\; 4x_2 - 2x_3 \;+\; \lambda_1(2x_1 - x_2 - x_3) \;+\; \lambda_2(x_1^2 + x_2^2 - 13). \end{aligned} \tag{9.4}$$

In the Lagrange multipliers method, the constrained extremum point x_0 is found by identifying the critical points of the Lagrangian $F(x,\lambda)$. For the method to work, the following necessary condition must be satisfied:

The gradient $\nabla g(x)$ has full rank at any point x where the constraint $g(x) = 0$ is satisfied, i.e.,

$$\mathit{rank}(\nabla g(x)) = m, \quad \forall\; x \in U \;\; \textit{such that}\;\; g(x) = 0. \tag{9.5}$$

Example (continued): In order to use the Lagrange multipliers method for our example, we first check that $\operatorname{rank}(\nabla g(x)) = 2$ for any $x \in \mathbb{R}^3$ such that $g(x) = 0$. Recall that the function $g(x)$ is given by (9.3). Then,

$$\nabla g(x) \;=\; \begin{pmatrix} 2 & -1 & -1 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}.$$

Note that $\text{rank}(\nabla g(x)) = 1$ if and only if $x_1 = x_2 = 0$. Also, $g(x) = 0$ if and only if $2x_1 - x_2 - x_3 = 0$ and $x_1^2 + x_2^2 = 13$. However, if $x_1^2 + x_2^2 = 13$, then it is not possible to have $x_1 = x_2 = 0$. We conclude that, if $g(x) = 0$, then $\text{rank}(\nabla g(x)) = 2$.

The gradient[1] of $F(x, \lambda)$ with respect to both x and λ will be denoted by $\nabla_{(x,\lambda)} F(x, \lambda)$ and is the following row vector:

$$\nabla_{(x,\lambda)} F(x,\lambda) \;=\; \left(\begin{array}{cc} \nabla_x F(x,\lambda) & \nabla_\lambda F(x,\lambda) \end{array} \right); \tag{9.6}$$

cf. (1.38). It is easy to see that

$$\frac{\partial F}{\partial x_j}(x,\lambda) \;=\; \frac{\partial f}{\partial x_j}(x) \;+\; \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(x), \quad \forall\ j = 1:n; \tag{9.7}$$

$$\frac{\partial F}{\partial \lambda_i}(x,\lambda) \;=\; g_i(x), \quad \forall\ i = 1:m. \tag{9.8}$$

Denote by $\nabla f(x)$ and $\nabla g(x)$ the gradients of $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}^m$, i.e.,

$$\nabla f(x) \;=\; \left(\frac{\partial f}{\partial x_1}(x,\lambda) \;\; \ldots \;\; \frac{\partial f}{\partial x_n}(x,\lambda) \right); \tag{9.9}$$

$$\nabla g(x) \;=\; \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \frac{\partial g_1}{\partial x_2}(x) & \ldots & \frac{\partial g_1}{\partial x_n}(x) \\ \frac{\partial g_2}{\partial x_1}(x) & \frac{\partial g_2}{\partial x_2}(x) & \ldots & \frac{\partial g_2}{\partial x_n}(x) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_m}{\partial x_1}(x) & \frac{\partial g_m}{\partial x_2}(x) & \ldots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix}; \tag{9.10}$$

cf. (1.38) and (1.40), respectively.

Then, from (9.7–9.10), it follows that

$$\nabla_x F(x,\lambda) \;=\; \left(\frac{\partial F}{\partial x_1}(x,\lambda) \;\ldots\; \frac{\partial F}{\partial x_n}(x,\lambda) \right) = \nabla f(x) + \lambda^t(\nabla g(x)); \tag{9.11}$$

$$\nabla_\lambda F(x,\lambda) \;=\; \left(\frac{\partial F}{\partial \lambda_1}(x,\lambda) \;\ldots\; \frac{\partial F}{\partial \lambda_m}(x,\lambda) \right) = (g(x))^t. \tag{9.12}$$

From (9.6), (9.11), and (9.12), we conclude that

$$\nabla_{(x,\lambda)} F(x,\lambda) \;=\; \left(\begin{array}{cc} \nabla f(x) + \lambda^t(\nabla g(x)) & (g(x))^t \end{array} \right). \tag{9.13}$$

The following theorem gives necessary conditions for a point $x_0 \in U$ to be a constrained extremum point for $f(x)$. Its proof involves the Inverse Function Theorem and is beyond the scope of this book.

---

[1] In this section, we use the notation ∇F, instead of DF, for the gradient of F.

Theorem 9.1. *Assume that the constraint function $g(x)$ satisfies the condition (9.5). If $x_0 \in U$ is a constrained extremum point of $f(x)$ with respect to the constraint $g(x) = 0$, then there exists a Lagrange multiplier $\lambda_0 \in \mathbb{R}^m$ such that the point (x_0, λ_0) is a critical point for the Lagrangian function $F(x, \lambda)$, i.e.,*

$$\nabla_{(x,\lambda)} \; F(x_0, \lambda_0) \;=\; 0. \tag{9.14}$$

We note that $\nabla_{(x,\lambda)} \; F(x, \lambda)$ is a function from $\mathbb{R}^{m+n}$ into $\mathbb{R}^{m+n}$. Thus, solving (9.14) to find the critical points of $F(x, \lambda)$ requires, in general, using the $N-$dimensional Newton's method for solving nonlinear equations; see section 5.2.1 for details. Interestingly enough, for some financial applications such as finding minimum variance portfolios, problem (9.14) is a linear system which can be solved without using Newton's method; see section 9.3 for details.

Example (continued): Since we already checked that the condition (9.5) is satisfied for this example, we proceed to find the critical points of the Lagrangian $F(x, \lambda)$. Using formula (9.4) for $F(x, \lambda)$, it is easy to see that

$$\nabla_{(x,\lambda)} F(x, \lambda) \;=\; \begin{pmatrix} 2\lambda_1 + 2\lambda_2 x_1 \\ 4 - \lambda_1 + 2\lambda_2 x_2 \\ -2 - \lambda_1 \\ 2x_1 - x_2 - x_3 \\ x_1^2 + x_2^2 - 13 \end{pmatrix}^t .$$

Let $x_0 = (x_{0,1}, x_{0,2}, x_{0,3})$ and $\lambda_0 = (\lambda_{0,1}, \lambda_{0,2})$. Then, solving $\nabla_{(x,\lambda)} F(x_0, \lambda_0) = 0$ is equivalent to solving the following system:

$$\begin{cases} 2\lambda_{0,1} + 2\lambda_{0,2} \; x_{0,1} \;=\; 0 \\ 4 - \lambda_{0,1} + 2\lambda_{0,2} \; x_{0,2} \;=\; 0 \\ -2 - \lambda_{0,1} \;=\; 0 \\ 2x_{0,1} - x_{0,2} - x_{0,3} \;=\; 0 \\ x_{0,1}^2 + x_{0,2}^2 - 13 \;=\; 0 \end{cases} \tag{9.15}$$

From the third equation of (9.15), we find that $\lambda_{0,1} = -2$. Then, the system (9.15) can be written as

$$\begin{cases} \lambda_{0,2} \; x_{0,1} \;=\; 2 \\ \lambda_{0,2} \; x_{0,2} \;=\; -3 \\ x_{0,3} \;=\; 2x_{0,1} - x_{0,2} \\ x_{0,1}^2 + x_{0,2}^2 \;=\; 13 \end{cases} \tag{9.16}$$

Since $\lambda_{0,2} \neq 0$, we find from (9.16) that

$$x_{0,1} = \frac{2}{\lambda_{0,2}}; \quad x_{0,2} = \frac{-3}{\lambda_{0,2}}; \quad x_{0,3} = \frac{7}{\lambda_{0,2}}; \quad x_{0,1}^2 + x_{0,2}^2 = \frac{13}{\lambda_{0,2}^2} = 13.$$

Thus, $\lambda_{0,2}^2 = 1$ and the system (9.15) has two solutions, one corresponding to $\lambda_{0,2} = 1$, and another one corresponding to $\lambda_{0,2} = -1$, as follows:

$$x_{0,1} = 2; \;\; x_{0,2} = -3; \;\; x_{0,3} = 7; \;\; \lambda_{0,1} = -2; \;\; \lambda_{0,2} = 1 \tag{9.17}$$

and

$$x_{0,1} = -2; \; x_{0,2} = 3; \; x_{0,3} = -7; \; \lambda_{0,1} = -2; \; \lambda_{0,2} = -1. \tag{9.18}$$

From (9.17) and (9.18), we conclude that the Lagrangian $F(x, \lambda)$ has the following two critical points:

$$x_0 \; = \; \begin{pmatrix} 2 \\ -3 \\ 7 \end{pmatrix}; \quad \lambda_0 \; = \; \begin{pmatrix} -2 \\ 1 \end{pmatrix} \tag{9.19}$$

and

$$x_0 \; = \; \begin{pmatrix} -2 \\ 3 \\ -7 \end{pmatrix}; \quad \lambda_0 \; = \; \begin{pmatrix} -2 \\ -1 \end{pmatrix}. \tag{9.20}$$

Finding sufficient conditions for a critical point (x_0, λ_0) of the Lagrangian $F(x, \lambda)$ to correspond to a constrained extremum point x_0 for $f(x)$ is somewhat more complicated (and rather rarely checked in practice).

Consider the function $F_0 : U \to \mathbb{R}$ given by

$$F_0(x) \; = \; F(x, \lambda_0) \; = \; f(x) \; + \; \lambda_0^t g(x).$$

Let $D^2 F_0(x_0)$ be the Hessian of $F_0(x)$ evaluated at the point x_0, i.e.,

$$D^2 F_0(x_0) \; = \; \begin{pmatrix} \frac{\partial^2 F_0}{\partial x_1^2}(x_0) & \frac{\partial^2 F_0}{\partial x_2 \partial x_1}(x_0) & \cdots & \frac{\partial^2 F_0}{\partial x_n \partial x_1}(x_0) \\ \frac{\partial^2 F_0}{\partial x_1 \partial x_2}(x_0) & \frac{\partial^2 F_0}{\partial x_2^2}(x_0) & \cdots & \frac{\partial^2 F_0}{\partial x_n \partial x_2}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F_0}{\partial x_1 \partial x_n}(x_0) & \frac{\partial^2 F_0}{\partial x_2 \partial x_n}(x_0) & \cdots & \frac{\partial^2 F_0}{\partial x_n^2}(x_0) \end{pmatrix}. \tag{9.21}$$

Note that $D^2 F_0(x_0)$ is an $n \times n$ matrix.

Let $q(v)$ be the quadratic form associated to the matrix $D^2 F_0(x_0)$, i.e.,

$$q(v) \; = \; v^t \; D^2 F_0(x_0) \; v \; = \; \sum_{1 \leq i,j \leq n} \frac{\partial^2 F_0}{\partial x_i \partial x_j}(x_0) v_i v_j, \tag{9.22}$$

where $v = (v_i)_{i=1:n}$.

We restrict our attention to the vectors v satisfying $\nabla g(x_0) \; v = 0$, i.e., to the vector space

$$V_0 \; = \; \{v \in \mathbb{R}^n \mid \nabla g(x_0) \; v = 0\}. \tag{9.23}$$

Note that $\nabla g(x_0)$ is a matrix with m rows and n columns, where $m < n$; cf. (9.10). If the condition (9.5) is satisfied, it follows that $\text{rank}(\nabla g(x_0)) = m$, i.e., the matrix $\nabla g(x_0)$ has m linearly independent columns. Assume, without losing any generality, that the first m columns of $\nabla g(x_0)$ are linearly independent.

By solving the linear system $\nabla g(x_0)\ v = 0$, we obtain that the entries v_1, v_2, $\ldots$, v_m of the vector v can be written as linear combinations of v_{m+1}, v_{m+2}, $\ldots$, v_n, the other $n-m$ entries of v. Let

$$v_{red} \ = \ \begin{pmatrix} v_{m+1} \\ \vdots \\ v_n \end{pmatrix}.$$

Then, by restricting $q(v)$ to the vector space V_0, we can write $q(v)$ as a quadratic form depending only on the entries of the vector v_{red}, i.e.,

$$q(v) \ = \ q_{red}(v_{red}) \ = \ \sum_{m+1\le i,j\le n} q_{red}(i,j) v_i v_j, \quad \forall\ v \in V_0. \tag{9.24}$$

Example (continued): We compute the reduced quadratic forms corresponding to the critical points of the Lagrangian function from our example. Recall that there are two such critical points, given by (9.19) and (9.20), respectively.

• The first critical point of $F(x,\lambda)$ is $x_0 = (2,-3,7)$ and $\lambda_0 = (-2,1)$; cf. (9.19). The function $F_0(x) = f(x) + \lambda_0^t g(x)$ is equal to

$$\begin{aligned} F_0(x) \ &= \ 4x_2 - 2x_3 + (-2)\cdot(2x_1 - x_2 - x_3) + 1\cdot(x_1^2 + x_2^2 - 13) \\ &= \ x_1^2 + x_2^2 - 4x_1 + 6x_2 - 13. \end{aligned} \tag{9.25}$$

From (9.25) and (9.26), we find that

$$D^2F_0(x_0) \ = \ \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \nabla g(x_0) \ = \ \begin{pmatrix} 2 & -1 & -1 \\ 4 & -6 & 0 \end{pmatrix}.$$

If $v = (v_i)_{i=1:3}$, then

$$q(v) \ = \ v^t\ D^2F_0(x_0)\ v \ = \ 2v_1^2 + 2v_2^2.$$

Recall that the gradient of the constraint function $g(x)$ is

$$\nabla g(x) \ = \ \begin{pmatrix} 2 & -1 & -1 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}. \tag{9.26}$$

Then, $\nabla g(x_0) = \begin{pmatrix} 2 & -1 & -1 \\ 4 & -6 & 0 \end{pmatrix}$, and the condition $\nabla g(x_0)v = 0$ is equivalent to

$$\left\{ \begin{array}{rcl} 2v_1 - v_2 - v_3 &=& 0 \\ 4v_1 - 6v_2 &=& 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} v_3 &=& 2v_1 - v_2 \\ v_1 &=& \frac{3}{2}v_2 \end{array} \right. \iff \left\{ \begin{array}{rcl} v_3 &=& 2v_2; \\ v_1 &=& \frac{3}{2}v_2. \end{array} \right.$$

Let $v_{red} = v_2$. The reduced quadratic form $q_{red}(v_{red})$ is

$$q_{red}(v_{red}) \ = \ 2v_1^2 + 2v_2^2 \ = \ \frac{13}{2}v_2^2. \tag{9.27}$$

• The second critical point of $F(x,\lambda)$ is $x_0 = (-2,3,-7)$ and $\lambda_0 = (-2,-1)$; cf. (9.20). The function $F_0(x) = f(x) + \lambda_0^t g(x)$ is equal to

$$F_0(x) \;=\; -\,x_1^2 - x_2^2 - 4x_1 + 6x_2 + 13. \tag{9.28}$$

From (9.28) and (9.26), we find that

$$D^2F_0(x_0) \;=\; \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \nabla g(x_0) \;=\; \begin{pmatrix} 2 & -1 & -1 \\ -4 & 6 & 0 \end{pmatrix}.$$

If $v = (v_i)_{i=1:3}$, then

$$q(v) \;=\; v^t\; D^2F_0(x_0)\; v \;=\; -\,2v_1^2 - 2v_2^2.$$

Then, $\nabla g(x_0) = \begin{pmatrix} 2 & -1 & -1 \\ -4 & 6 & 0 \end{pmatrix}$ and the condition $\nabla g(x_0)v = 0$ is equivalent to

$$\left\{ \begin{array}{rcl} 2v_1 - v_2 - v_3 & = & 0 \\ -4v_1 + 6v_2 & = & 0 \end{array} \right. \Longleftrightarrow \left\{ \begin{array}{rcl} v_3 & = & 2v_1 - v_2 \\ v_1 & = & \frac{3}{2}v_2 \end{array} \right. \Longleftrightarrow \left\{ \begin{array}{rcl} v_3 & = & 2v_2; \\ v_1 & = & \frac{3}{2}v_2. \end{array} \right.$$

Let $v_{red} = v_2$. The reduced quadratic form $q_{red}(v_{red})$ is

$$q_{red}(v_{red}) \;=\; -\,2v_1^2 - 2v_2^2 \;=\; -\,\frac{13}{2}v_2^2. \quad \square \tag{9.29}$$

Whether the point x_0 is a constrained extremum for $f(x)$ will depend on the nonzero quadratic form q_{red} being either positive semidefinite, i.e.,

$$q_{red}(v_{red}) \;\geq\; 0, \quad \forall\; v_{red} \in \mathbb{R}^{n-m},$$

or negative semidefinite, i.e.,

$$q_{red}(v_{red}) \;\leq\; 0, \quad \forall\; v_{red} \in \mathbb{R}^{n-m}.$$

Theorem 9.2. *Assume that the constraint function $g(x)$ satisfies condition (9.5). Let $x_0 \in U \subset \mathbb{R}^n$ and $\lambda_0 \in \mathbb{R}^m$ such that the point (x_0, λ_0) is a critical point for the Lagrangian function $F(x,\lambda) = f(x) + \lambda^t g(x)$.*

If the reduced quadratic form (9.24) corresponding to the point (x_0,λ_0) is positive semidefinite, then x_0 is a constrained minimum for $f(x)$ with respect to the constraint $g(x) = 0$.

If the reduced quadratic form (9.24) corresponding to the point (x_0,λ_0) is negative semidefinite, then x_0 is a constrained maximum for $f(x)$ with respect to the constraint $g(x) = 0$.

If the reduced quadratic form (9.24) corresponding to (x_0,λ_0) is nonzero, and it is not positive semidefinite nor negative semidefinite, then x_0 is not a constrained extremum point for $f(x)$ with respect to the constraint $g(x) = 0$.

Example (continued): We use Theorem 9.2 to classify the critical points of the Lagrangian function corresponding to our example.

Recall from (9.27) that the reduced quadratic form corresponding to the critical point $x_0 = (2, -3, 7)$ and $\lambda_0 = (-2, 1)$ of $F(x, \lambda)$ is $q_{red}(v_{red}) = \frac{13}{2}v_2^2$, which is positive semidefinite for any $x \in \mathbb{R}^3$. Using Theorem 9.2, we conclude that the point $(2, -3, 7)$ is a minimum point for $f(x)$. By direct computation, we find that $f(2, -3, 7) = -26$.

Recall from (9.29) that the reduced quadratic form corresponding to the critical point $x_0 = (-2, 3, -7)$ and $\lambda_0 = (-2, -1)$ of $F(x, \lambda)$ is $q_{red}(v_{red}) = -\frac{13}{2}v_2^2$, which is negative semidefinite for any $x \in \mathbb{R}^3$. Using Theorem 9.2, we conclude that the point $(-2, 3, -7)$ is a maximum point for $f(x)$. By direct computation, we find that $f(-2, 3, -7) = 26$. □

It is important to note that the solution to the constrained optimization problem presented in Theorem 9.2 is significantly simpler, and does not involve computing the reduced quadratic form (9.24), if the matrix $D^2F_0(x_0)$ is either positive definite or negative definite.

Corollary 9.1. *Assume that the constraint function $g(x)$ satisfies condition (9.5). Let $x_0 \in U \subset \mathbb{R}^n$ and $\lambda_0 \in \mathbb{R}^m$ such that the point (x_0, λ_0) is a critical point for the Lagrangian function $F(x, \lambda) = f(x) + \lambda^t g(x)$. Let $F_0(x) = f(x) + \lambda_0^t g(x)$, and let $D^2F_0(x_0)$ be the Hessian of F_0 evaluated at the point x_0.*

If the matrix $D^2F_0(x_0)$ is positive definite, i.e., if all the eigenvalues of the matrix $D^2F_0(x_0)$ are strictly greater than 0, then x_0 is a constrained minimum for $f(x)$ with respect to the constraint $g(x) = 0$.

If the matrix $D^2F_0(x_0)$ is negative definite, i.e., if all the eigenvalues of the matrix $D^2F_0(x_0)$ are strictly less than 0, then x_0 is a constrained maximum for $f(x)$ with respect to the constraint $g(x) = 0$.

The result of Corollary 9.1 will be used in sections 9.3 and 9.4 for details to find minimum variance portfolios and maximum return portfolios, respectively.

is not straightforward.

Summarizing, the steps required to solve a constrained optimization problem using Lagrange multipliers are:

Step 1: Check that $\text{rank}(\nabla g(x)) = m$ for all x such that $g(x) = 0$.

Step 2: Find $(x_0, \lambda_0) \in U \times \mathbb{R}^m$ such that $\nabla_{(x,\lambda)} F(x_0, \lambda_0) = 0$.

Step 3.1: Compute $q(v) = v^t \ D^2F_0(x_0) \ v$, where $F_0(x) = f(x) + \lambda_0^t g(x)$.

Step 3.2: Compute $q_{red}(v_{red})$ by restricting $q(v)$ to the vectors v satisfying the condition $\nabla g(x_0) \ v = 0$. Decide whether $q_{red}(v_{red})$ is positive semidefinite or negative semidefinite.

Step 4: Use Theorem 9.2 to decide whether x_0 is a constrained minimum point or a constrained maximum point.

Note: If the matrix $D^2F_0(x_0)$ is either positive semidefinite or negative semidefinite, skip Step 3.2 and go from Step 3.1 to the following version of Step 4:

Step 4: Use Corollary 9.1 to decide whether x_0 is a constrained minimum point or a constrained maximum point.

Several examples of solving constrained optimization problems using Lagrange multipliers are given in section 9.1.1; the steps above are outlined for each example. Applications of the Lagrange multipliers method to portfolio optimization problems are presented in sections 9.3 and 9.4.

9.1.1 Examples

The first example below illustrates the fact that, if condition (9.5) is not satisfied, then Theorem 9.1 may not hold.

Example: Find the minimum value of $x_1^2 + x_1 + x_2^2$, subject to the constraint $(x_1 - x_2)^2 = 0$.

Answer: This problem can be solved without using Lagrange multipliers as follows: If $(x_1 - x_2)^2 = 0$, then $x_1 = x_2$. Then, the function to minimize becomes

$$x_1^2 + x_1 + x_2^2 \;=\; 2x_1^2 + x_1 \;=\; 2\left(x_1 + \frac{1}{4}\right)^2 - \frac{1}{8},$$

which achieves its minimum when $x_1 = -\frac{1}{4}$. We conclude that there exists a unique constrained minimum point $x_1 = x_2 = -\frac{1}{4}$, and that the corresponding minimum value is $-\frac{1}{8}$.

Although we solved the problem directly, we attempt to find an alternative solution using the Lagrange multipliers method. According to the framework outlined in section 9.1, we want to find the minimum of the function

$$f(x_1, x_2) \;=\; x_1^2 + x_1 + x_2^2$$

for $(x_1, x_2) \in \mathbb{R}^2$, such that the constraint $g(x_1, x_2) = 0$ is satisfied, where

$$g(x_1, x_2) \;=\; (x_1 - x_2)^2.$$

By definition,

$$\nabla g(x) \;=\; \begin{pmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \end{pmatrix} \;=\; \begin{pmatrix} 2(x_1 - x_2) & -2(x_1 - x_2) \end{pmatrix}.$$

Note that $g(x) = 0$ if and only if $x_1 = x_2$. Thus, $\nabla g(x) = (0\ 0)$ (and therefore $\text{rank}(\nabla g(x)) = 0$) for all x such that $g(x) = 0$. We conclude that condition (9.5) is not satisfied at any point x such that $g(x) = 0$.

Since the problem has one constraint, we only have one Lagrange multiplier, which we denote by $\lambda \in \mathbb{R}$. From (9.2), it follows that the Lagrangian is

$$F(x_1, x_2, \lambda) \;=\; f(x_1, x_2) \;+\; \lambda g(x_1, x_2) \;=\; x_1^2 + x_1 + x_2^2 \;+\; \lambda(x_1 - x_2)^2.$$

The gradient $\nabla_{(x,\lambda)} F(x, \lambda)$ is computed as in (9.13) and has the form

$$\nabla_{(x,\lambda)}\; F(x, \lambda) = \begin{pmatrix} 2x_1 + 1 + 2\lambda(x_1 - x_2) & 2x_2 - 2\lambda(x_1 - x_2) & (x_1 - x_2)^2 \end{pmatrix}.$$

Finding the critical points for $F(x, \lambda)$ requires solving $\nabla_{(x,\lambda)}\; F(x, \lambda) = 0$, which is equivalent to the following system of equations:

$$\begin{cases} 2x_1 + 1 + 2\lambda(x_1 - x_2) &= 0; \\ 2x_2 - 2\lambda(x_1 - x_2) &= 0; \\ (x_1 - x_2)^2 &= 0. \end{cases}$$

This system does not have a solution: from the third equation, we obtain that $x_1 = x_2$. Then, from the second equation, we find that $x_2 = 0$, which implies that $x_1 = 0$. Substituting $x_1 = x_2 = 0$ into the first equation, we obtain $1 = 0$, which is a contradiction.

In other words, the Lagrangian $F(x, \lambda)$ has no critical points. However, we showed before that the point $(x_1, x_2) = \left(-\frac{1}{4}, -\frac{1}{4}\right)$ is a constrained minimum point for $f(x)$ given the constraint $g(x) = 0$. The reason Theorem 9.1 does not apply in this case is that condition (9.5), which was required in order for Theorem 9.1 to hold, is not satisfied. $\square$

Example: Find the positive numbers x_1, x_2, x_3 such that $x_1 x_2 x_3 = 1$ and $x_1 x_2 + x_2 x_3 + x_3 x_1$ is minimized.

Answer: We first reformulate the problem as a constrained optimization problem. Let $U = \prod_{i=1:3}(0, \infty) \subset \mathbb{R}^3$ and let $x = (x_1, x_2, x_3) \in U$. The functions $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}$ are defined as

$$f(x) \;=\; x_1 x_2 + x_2 x_3 + x_3 x_1; \quad g(x) \;=\; x_1 x_2 x_3 - 1.$$

We want to minimize $f(x)$ over the set U subject to the constraint $g(x) = 0$.

<u>Step 1</u>: Check that $\text{rank}(\nabla g(x)) = 1$ for any x such that $g(x) = 0$.
Let $x = (x_1, x_2, x_3) \in U$. It is easy to see that

$$\nabla g(x) \;=\; \begin{pmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \end{pmatrix}. \tag{9.30}$$

Note that $\nabla g(x) \neq 0$, since $x_i > 0$, $i = 1 : 3$. Therefore, $\text{rank}(\nabla g(x)) = 1$ for all $x \in U$, and condition (9.5) is satisfied.

<u>Step 2</u>: Find (x_0, λ_0) such that $\nabla_{(x,\lambda)} F(x_0, \lambda_0) = 0$.
The Lagrangian associated to this problem is

$$F(x, \lambda) \;=\; x_1 x_2 + x_2 x_3 + x_3 x_1 \;+\; \lambda\,(x_1 x_2 x_3 - 1), \tag{9.31}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier.

Let $x_0 = (x_{0,1}, x_{0,2}, x_{0,3})$. From (9.31), we find that $\nabla_{(x,\lambda)} F(x_0, \lambda_0) = 0$ is equivalent to the following system:

$$\begin{cases} x_{0,2} + x_{0,3} + \lambda_0 \; x_{0,2} \; x_{0,3} & = & 0; \\ x_{0,1} + x_{0,3} + \lambda_0 \; x_{0,1} \; x_{0,3} & = & 0; \\ x_{0,1} + x_{0,2} + \lambda_0 \; x_{0,1} \; x_{0,2} & = & 0; \\ x_{0,1} \; x_{0,2} \; x_{0,3} & = & 1. \end{cases}$$

By multiplying the first three equations by $x_{0,1}$, $x_{0,2}$, and $x_{0,3}$, respectively, and using the fact that $x_{0,1}x_{0,2}x_{0,3} = 1$, we obtain that

$$-\lambda \;=\; x_{0,1}\; x_{0,2} + x_{0,1}\; x_{0,3} \;=\; x_{0,1}\; x_{0,2} + x_{0,2}\; x_{0,3} \;=\; x_{0,1}\; x_{0,3} + x_{0,2}\; x_{0,3}.$$

Since $x_{0,i} \neq 0$, $i = 1:3$, we find that $x_{0,1} = x_{0,2} = x_{0,3}$. Since $x_{0,1}\; x_{0,2}\; x_{0,3} = 1$, we conclude that $x_{0,1} = x_{0,2} = x_{0,3} = 1$ and $\lambda_0 = -2$.

Step 3.1: Compute $q(v) = v^t \; D^2 F_0(x_0) \; v$.
Since $\lambda_0 = -2$, we find that $F_0(x) = f(x) + \lambda_0^t g(x)$ is given by

$$F_0(x) \;=\; x_1x_2 + x_2x_3 + x_3x_1 - 2x_1x_2x_3 + 2.$$

The Hessian of $F_0(x)$ is

$$D^2F_0(x_1, x_2, x_3) \;=\; \begin{pmatrix} 0 & 1-2x_3 & 1-2x_2 \\ 1-2x_3 & 0 & 1-2x_1 \\ 1-2x_2 & 1-2x_1 & 0 \end{pmatrix},$$

and therefore

$$D^2F_0(1,1,1) \;=\; \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

From (9.22), we find that the quadratic form $q(v)$ is

$$q(v) \;=\; v^t \; D^2F_0(1,1,1) \; v \;=\; -\, 2v_1v_2 - 2v_2v_3 - 2v_1v_3. \tag{9.32}$$

Step 3.2: Compute $q_{red}(v_{red})$.
We first solve formally the equation $\nabla g(1,1,1)\; v = 0$, where $v = (v_i)_{i=1:3}$ is an arbitrary vector. From (9.30), we find that $\nabla g(1,1,1) = (1\ 1\ 1)$, and

$$\nabla g(1,1,1)\; v \;=\; v_1 + v_2 + v_3 \;=\; 0.$$

By solving for v_1 in terms of v_2 and v_3 we obtain that $v_1 = -v_2 - v_3$. Let $v_{red} = \begin{pmatrix} v_2 \\ v_3 \end{pmatrix}$. We substitute $-v_2 - v_3$ for v_1 in (9.32), to obtain the reduced quadratic form $q_{red}(v_{red})$, i.e.,

$$q_{red}(v_{red}) \;=\; 2v_2^2 + 2v_2v_3 + 2v_3^2 \;=\; v_2^2 + v_3^2 + (v_2 + v_3)^2.$$

Therefore, $q_{red}(v_{red}) > 0$ for all $(v_2, v_3) \neq (0,0)$, which means that q_{red} is a positive definite form.

Step 4: From Theorem 9.2, we conclude that the point $x_0 = (1,1,1)$ is a constrained minimum for the function $f(x) = x_1x_2 + x_2x_3 + x_3x_1$, with $x_1, x_2, x_3 > 0$, subject to the constraint $x_1x_2x_3 = 1$. $\square$