Appendix D

Optimization in finite-dimensional vector space

This appendix introduces the concepts and nomenclature associated with the optimization problem – both the constrained and the unconstrained versions in a finite-dimensional vector space. The aim is to provide a characterization of the properties of the optimal solution for both the constrained and unconstrained minimization problem.

D.1 An optimization problem

Let Ω be a subset of $\mathbb{R}^n$, and let $\phi : \Omega \longrightarrow \mathbb{R}$ denote a real-valued function. For reasons that will become apparent soon, it is assumed that ϕ satisfies the following condition:

Condition C ϕ along with the following set of partial derivatives, namely, $\frac{\partial \phi}{\partial x_i}$, $\frac{\partial^2 \phi}{\partial x_i^2}$ and $\frac{\partial^2 \phi}{\partial x_i \partial x_j}$ for $1 \le i, j \le n$ and $i \neq j$ are **continuous** functions of $\mathbf{x} \in \Omega$. That is, ϕ belongs to the class of twice continuously differentiable functions.

A function $\phi(\mathbf{x})$ is said to attain a **relative** or **local minimum** at $\mathbf{x} = \mathbf{x}^*$ in Ω, if

$$\phi(\mathbf{x}^*) \le \phi(\mathbf{x}) \tag{D.1.1}$$

for all points $\mathbf{x}$ in a sufficiently small neighborhood $N_\epsilon(\mathbf{x}^*) = \{\mathbf{x} | \, \|\mathbf{x} - \mathbf{x}^*\| < \epsilon\}$. If strict inequality in (D.1.1) holds for all $N_\epsilon(\mathbf{x}^*)$ except $\mathbf{x}^*$, then $\mathbf{x}^*$ is called a *strict local minimum*. Likewise $\mathbf{x}^*$ is called the *absolute* or *global minimum* if (D.1.1) holds for all $\mathbf{x} \in \Omega$. If strict inequality in (D.1.1) holds for all $\mathbf{x} \in \Omega$ except $\mathbf{x}^*$, then $\mathbf{x}^*$ is known as the **strict global minimum**.

A version of the optimization problem may be stated as follows.

Problem P Given Ω and the function ϕ satisfying condition **C**, find the set of points in Ω where $\phi(\mathbf{x})$ attains minimum value.

If Ω is a proper subset of $\mathbb{R}^n$, then problem **P** is called a **constrained** minimization problem, otherwise, that is when $\Omega = \mathbb{R}^n$, it is called **unconstrained** minimization problem. Ω is often called the **feasible set**.

Example D.1.1 Let $\mathbf{x} \in \mathbb{R}^2$. An instance of the unconstrained minimization problem is to find the points at which

$$\phi(\mathbf{x}) = 16x_1^2 + 4x_1x_2 + 4x_2^2 - 7x_1 + 5x_2 + 6$$

is a minimum.

Example D.1.2 Let $\mathbf{x} \in \mathbb{R}^3$. An instance of the constrained optimization problem with **inequality** constraints is to find the minimum of

$$\phi(\mathbf{x}) = x_1x_2x_3$$

when Ω is given by

$$\Omega = \left\{\mathbf{x} | x_1^2 + x_2^2 + x_3^2 \le 4 \quad \text{and} \quad x_i > 0, \;\; 1 \le i \le 3\right\}.$$

Notice that in this case the feasible set consists of the points on and inside the sphere of radius 2 centered at the origin.

Example D.1.3 Let $\mathbf{x} \in \mathbb{R}^2$. An instance of the constrained minimization with **equality** constraints is to minimize $\phi(\mathbf{x}) = x_1^2 + x_2^2$, when Ω is defined by

$$\Omega = \{\mathbf{x} | x_1 + x_2 = 1\}.$$

In this case, the feasible set is a straight line.

One can easily visualize constrained problems having both equality and inequality constraints. In this book we will be concerned exclusively with minimization problems under equality constraints.

Remark D.1.1 When $\phi(\mathbf{x})$ and all the functions defining the constraint set Ω are linear functions of $\mathbf{x}$, then problem $\mathbf{P}$ is called a **linear programming** problem. If $\phi(\mathbf{x})$ is a **nonlinear** function, then problem $\mathbf{P}$ is called a **nonlinear programming** problem (Luenberger [1973]). Accordingly, the problem of determining the optimal initial conditions for deterministic models, which is one of the principal topics of this book, may be viewed as an exercise in nonlinear programming.

D.2 Condition for a minimum-unconstrained problem

In this section we consider the unconstrained minimization problem. In particular, we provide a characterization of the properties of the optimal solution.

To fix our ideas, first consider $\phi : \mathbb{R} \longrightarrow \mathbb{R}$, a real-valued function of a real variable satisfying the condition $\mathbf{C}$. For any increment Δx in x, we can expand $\phi(x + \Delta x)$ using the standard Taylor series (Appendix C) as

$$\phi(x + \Delta x) = \phi(x) + \frac{\mathrm{d}\phi}{\mathrm{d}x}\Delta x + \frac{1}{2}\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2}(\Delta x)^2 + \cdots \tag{D.2.1}$$

If $\Delta\phi(x)$ denotes the *increment* in ϕ resulting from a small increment Δx in x, we have

$$\begin{aligned}\Delta\phi(x) &= \phi(x+\Delta x) - \phi(x)\\ &\approx \frac{\mathrm{d}\phi}{\mathrm{d}x}(\Delta x) + \frac{1}{2}\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2}(\Delta x)^2\\ &= \delta\phi(x, \Delta x) + \frac{1}{2}\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2}(\Delta x)^2\end{aligned} \tag{D.2.2}$$

where

$$\delta\phi(x, \Delta x) = \frac{\mathrm{d}\phi}{\mathrm{d}x}(\Delta x) \tag{D.2.3}$$

is called the **first variation** in ϕ resulting from Δx (Appendix C). Equation (D.2.2) constitutes the basis from which all sorts of conditions characterizing the minima are obtained.

For small values of Δx, since $\delta(x, \Delta x)$ is the **dominant** term in (D.2.2), we can represent $\Delta\phi(x)$ to **first-order accuracy** as

$$\Delta\phi(x) \approx \delta(x, \Delta x) = \frac{\mathrm{d}\phi}{\mathrm{d}x}(\Delta x). \tag{D.2.4}$$

Now, consider a point x^* where $\phi(x)$ is a relative minimum. Then, by definition $\Delta\phi(x^*) \geq 0$. From (D.2.4), this requires that $\mathrm{d}\phi(x^*)/\mathrm{d}x(\Delta x) \geq 0$ for arbitrary (both positive and negative) Δx. Since $\mathrm{d}\phi(x^*)/\mathrm{d}x$ is a constant, this implies that

$$\frac{\mathrm{d}\phi(x^*)}{\mathrm{d}x} = 0 \tag{D.2.5}$$

and hence $\delta\phi(x^*, \Delta x) = 0$ at the minimum. Condition (D.2.5) is called a *first-order necessary condition* for minimum.

To obtain the **second-order necessary condition**, since $\mathrm{d}\phi(x^*)/\mathrm{d}x = 0$, from (D.2.2), it follows that

$$0 \leq \Delta\phi(x^*) = \phi(x^* + \Delta x) - \phi(\mathbf{x}) = \frac{1}{2}\frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2}(\Delta x)^2$$

which in turn implies that

$$\frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2} \geq 0. \tag{D.2.6}$$

Conversely, let x^* be a point such that

$$\frac{\mathrm{d}\phi(x^*)}{\mathrm{d}x} = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2} > 0.$$

Then, from (D.2.2), we obtain

$$\Delta\phi(x^*) = \phi(x^* + \Delta x) - \phi(\mathbf{x}^*) = \frac{1}{2}\frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2}(\Delta x)^2 > 0$$

which, by definition, implies that x^* is a relative minimum. We summarize these basic facts as follows:

Necessary Condition In order that $\phi(x)$ attains a minimum at x^*, it is necessary that

$$\frac{\mathrm{d}\phi(x^*)}{\mathrm{d}x} = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2} \geq 0. \tag{D.2.7}$$

Sufficient condition If x^* is such that

$$\frac{\mathrm{d}\phi(x^*)}{\mathrm{d}x} = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\phi(x^*)}{\mathrm{d}x^2} > 0, \tag{D.2.8}$$

then, $\phi(x)$ attains a relative minimum at x^*.

Remark D.2.1 Let $I = [a, b]$ be a closed interval in $\mathbb{R}$ and let $\phi : I \to I$. If ϕ is continuous, then there **exists** at least one point x^* in I at which ϕ attains a minimum. In other words, the *existence* of a minimum is guaranteed if the domain is closed interval and ϕ is continuous. In this case, the minimum occurs either at an *interior point* x^* satisfying the condition in (D.2.7) or else it occurs at an end point. If the left end point a is a minimum, then $\mathrm{d}\phi(a)/\mathrm{d}x \geq 0$ and if the right end point b is a minimum then $\mathrm{d}\phi(b)/\mathrm{d}x \leq 0$.

Remark D.2.2 Let $\phi : I \longrightarrow I$, where $I = [a, b]$. If ϕ is **strictly convex** (refer to Appendix C), then ϕ has a **unique** minimum at $x^* \in I$. In other words, for a strictly convex function, the local minimum is indeed a global minimum as well.

Remark D.2.3 In the above discussion we have assumed that ϕ and its first two derivatives are continuous. Indeed, the notion of a minimum also carries over to functions not necessarily continuous, but their treatment is beyond our scope. Most of the minimization problems of interest to us in this book involve quadratic functions which will fit the framework developed herein.

We now extend the above development to the **multivariate** case namely $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$. That is, ϕ is a scalar function of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$. Let $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)^{\mathrm{T}}$ denote a vector of infinitesimal change in $\mathbf{x}$. Again, using the Taylor series expansion (Appendix C), the resulting scalar increment $\Delta\phi(x)$ in $\phi(\mathbf{x})$ is given by

$$\begin{aligned} \Delta\phi(x) &= \phi(\mathbf{x} + \Delta\mathbf{x}) - \phi(\mathbf{x}) \\ &= (\Delta\mathbf{x})^{\mathrm{T}}\nabla\phi(\mathbf{x}) + \tfrac{1}{2}(\Delta\mathbf{x})^{\mathrm{T}}\nabla^2\phi(\mathbf{x})\Delta\mathbf{x} \end{aligned} \tag{D.2.9}$$

where $\nabla\phi(\mathbf{x})$ is the **gradient** of $\phi(\mathbf{x})$ and $(\Delta\mathbf{x})^{\mathrm{T}}\nabla\phi(\mathbf{x}) = \delta\phi(\mathbf{x}, \Delta\mathbf{x})$ is called the **first variation** in ϕ that dominates the right-hand side of (D.2.9) when $\|\Delta\mathbf{x}\|$, the norm of the vector $\Delta\mathbf{x}$ (Appendix A) is small. $\nabla^2\phi(\mathbf{x})$ is the **Hessian** of $\phi(\mathbf{x})$ (Appendix C).

The **direction cosines** α_i, $1 \le i \le n$ of $\Delta\mathbf{x}$ are given by (Appendix A)

$$\alpha_i = \frac{\Delta x_i}{\|\Delta\mathbf{x}\|},\ 1 \le i \le n,$$

and let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\mathrm{T}}$. If $\mathbf{x}^*$ is a **local minimum**, then $\Delta\phi(\mathbf{x}^*) \ge 0$ from the definition. Combining this with the first-order representation of $\Delta\phi(\mathbf{x})$ using (D.2.9), we require that

$$\frac{\Delta\phi(\mathbf{x}^*)}{\|\Delta\mathbf{x}\|} = \frac{\delta\phi(\mathbf{x}^*, \Delta\mathbf{x})}{\|\Delta\mathbf{x}\|} = \alpha^{\mathrm{T}}\nabla\phi_{\ge}(\mathbf{x})0 \tag{D.2.10}$$

for *every* direction vector α. Clearly, this can happen only if

$$\nabla\phi(\mathbf{x}) = 0 \tag{D.2.11}$$

which is a **first-order necessary** condition for a minimum of ϕ. Notice again at the minimum $\delta\phi(\mathbf{x}^*, \Delta\mathbf{x}) = 0$. Using (D.2.11) in (D.2.9), it follows that

$$\Delta\phi(x^*) = \frac{\|\Delta\mathbf{x}\|^2}{2}\alpha^{\mathrm{T}}[\nabla^2\phi(\mathbf{x}^*)]\alpha \ge 0 \tag{D.2.12}$$

which, in turn, requires

$$\alpha^{\mathrm{T}}[\nabla^2\phi(\mathbf{x}^*)]\alpha \ge 0 \tag{D.2.13}$$

for all α. (D.2.13) dictates that the Hessian $\nabla^2\phi(\mathbf{x}^*)$ must be **nonnegative definite at the minimum.**

Conversely, let $\mathbf{x}^*$ be such that

$$\nabla\phi\,(\mathbf{x}) = 0 \text{ and } \nabla^2\phi(\mathbf{x}^*) \text{ is positive definite.} \tag{D.2.14}$$

Then,

$$\begin{aligned}\Delta\phi(\mathbf{x}^*) &= \phi(\mathbf{x}^* + \Delta\mathbf{x}) - \phi(\mathbf{x}^*)\\ &= (\Delta\mathbf{x})^{\mathrm{T}}\nabla\phi_{+}(\mathbf{x})\tfrac{1}{2}(\Delta\mathbf{x})^{\mathrm{T}}[\nabla^2\phi(\mathbf{x}^*)](\Delta\mathbf{x}) > 0.\end{aligned}$$

Hence, $\mathbf{x}^*$ is a minimum. Summarizing, we obtain the following:

Necessary condition In order that $\phi(\mathbf{x})$ is a minimum at $\mathbf{x}^*$, it is necessary that $\nabla\phi = (\mathbf{x})0$ and $\nabla^2\phi(\mathbf{x}^*)$ is non-negative definite.

Sufficiency condition If $\mathbf{x}^*$ is such that $\nabla\phi = (\mathbf{x})0$ and $\nabla^2\phi(\mathbf{x}^*)$ is positive definite, then $\phi(\mathbf{x})$ attains a local minimum at $\mathbf{x}^*$.

Multivariate analogs of the Remarks D.2.1 through D.2.3 carry over to $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$. The following example illustrates these concepts.

Example D.2.1 Considering the $\phi(\mathbf{x})$ in Example D.1.1., it can be verified that

$$\nabla\phi(\mathbf{x}) = \begin{pmatrix} 32x_1 + 4x_2 - 7 \\ 4x_1 + 8x_2 + 5 \end{pmatrix} = 2\begin{bmatrix} 16 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -7 \\ 5 \end{bmatrix}$$

and

$$\nabla^2 \phi(\mathbf{x}) = 2 \begin{bmatrix} 16 & 2 \\ 2 & 4 \end{bmatrix}.$$

Solving $\nabla \phi = (\mathbf{x})0$, we obtain $\mathbf{x}^* = (0.3167, -0.7833)^{\mathrm{T}}$. Since $\mathbf{H}(\mathbf{x}^*)$ is positive definite, indeed $\mathbf{x}^*$ is a strict global (why?) minimum.

In the following section, we move on to characterizing the properties of a minimum when there are equality constraints.

D.3 Lagrangian method: constrained problem

The problem is to minimize $\phi(\mathbf{x})$ when Ω is specified by a collection of k relations as follows:

$$f_i(\mathbf{x}) = 0, \ \ 1 \le i \le k \tag{D.3.1}$$

where we assume that each $f_i(\mathbf{x})$ satisfies the Condition **C** in the same manner as the function $\phi(\mathbf{x})$. Each of the constraint equations, in principle, defines a surface in $\mathbb{R}^n$. Clearly Ω consists of all points common to these k surfaces, the set of points formed by the intersection of these surfaces. A straightforward approach to solving this problem is elimination of k variables from $\phi(\mathbf{x})$ using the constraint equations and to solve the resulting unconstrained problem in $(n - k)$ variables.

We illustrate this idea in the following example:

Example D.3.1 Let $\phi(\mathbf{x}) = x_1^2 + x_2^2$ and $f_1(\mathbf{x}) = x_1 + x_2 - 1 = 0$. Eliminating x_2 in $\phi(\mathbf{x})$ it can be verified that $\phi(\mathbf{x}) = 2x_1^2 - 2x_1 + 1$ which attains its minimum at $x_1^* = x_2^* = 1/2$.

Implicit in the above discussion is the notion of **degrees of freedom**. If there are no constraints, the number of independent variables in $\phi(x_1, x_2, \ldots, x_n)$ is defined as the number of the degrees of freedom. This implies that the minimum could, in principle, be anywhere in the Euclidean space, $\mathbb{R}^n$. The addition of a constraint narrows the search region – the minimum must lie along the surface defined by the constraint. That is, each constraint reduces the degree of freedom by unity. Thus, if there are k constraints defining Ω, then the net degree of freedom is $(n - k)$. In example (D.3.1), the minimum lies along the line $x_1 + x_2 = 1$ in $\mathbb{R}^2$, and the degree of freedom is 1.

Only in exceptionally simple cases such as in Example D.3.1, a constrained problem can be transformed into an unconstrained problem by elimination. In general, a methodology is needed that circumvents the solution by elimination of variables. Such a method was developed by Lagrange and has come to be known as the **Lagrangian** method.

Lagrangian method: one equality constraint

Consider first the problem of minimizing $\phi(\mathbf{x})$ when Ω is defined by only one constraint equation

$$f(\mathbf{x}) = 0. \tag{D.3.2}$$

Let $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)^{\mathrm{T}}$ and $\Delta\mathbf{x}/\|\Delta\mathbf{x}\| = \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\mathrm{T}}$ be such that

$$f(\mathbf{x} + \Delta\mathbf{x}) = 0. \tag{D.3.3}$$

Then, in view of (D.3.2) and (D.3.3), the first variation of f, δf is

$$\delta f = f(\mathbf{x} + \Delta\mathbf{x}) - f(\mathbf{x}) = (\Delta\mathbf{x})^{\mathrm{T}}\nabla f(\mathbf{x}) = \|\Delta\mathbf{x}\|\,(\alpha)^{\mathrm{T}}\nabla f = 0. \tag{D.3.4}$$

In other words, (D.3.4) defines the set of all **feasible** perturbations of $\mathbf{x}$ satisfying the condition (D.3.2). A number of observations are in order. Notice first that the direction cosine α corresponding to the perturbation vector $\Delta\mathbf{x}$ is orthogonal to ∇f. Since ∇f is normal to the surface $f(\mathbf{x}) = 0$, it follows that the direction cosine α lies in the plane that is tangential to $f(\mathbf{x})$ at the point $\mathbf{x}$. Secondly, equation (D.3.4) implies that not all of the components of the vector $\Delta\mathbf{x}$ are independent. That is, we can express one of the components, say Δx_n (assuming $\partial f/\partial x_n \neq 0$) in terms of the other components as

$$\Delta x_n = -\left(\frac{\partial f}{\partial x_n}\right)^{-1}\left[\left(\frac{\partial f}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial f}{\partial x_2}\right)\Delta x_2 + \cdots + \left(\frac{\partial f}{\partial x_{n-1}}\right)\Delta x_{n-1}\right]. \tag{D.3.5}$$

The induced first variation in $\phi(\mathbf{x})$ is given by

$$\delta\phi = \phi(\mathbf{x} + \Delta\mathbf{x}) - \phi(\mathbf{x}) = (\Delta\mathbf{x})^{\mathrm{T}}\nabla\phi(\mathbf{x}). \tag{D.3.6}$$

If $\mathbf{x}$ is a relative minimum, then $\delta\phi \geq 0$ for all direction cosine α satisfying (D.3.4). In view of this constraint we cannot automatically require $\nabla\phi = 0$ as in the unconstrained case. To find the condition for a minimum of ϕ, first consider the sum

$$\delta\phi + \lambda\delta f \tag{D.3.7}$$

for some yet to be determined scalar multiplier λ. At the relative minimum of $\phi(\mathbf{x})$, clearly (from D.3.4 and $\delta\phi \geq 0$ at the minimum)

$$\delta\phi + \lambda\delta f \geq 0 \tag{D.3.8}$$

which becomes

$$\left(\frac{\partial\phi}{\partial x_1} + \lambda\frac{\partial f}{\partial x_1}\right)\Delta x_1 + \left(\frac{\partial\phi}{\partial x_2} + \lambda\frac{\partial f}{\partial x_2}\right)\Delta x_2 + \cdots + \left(\frac{\partial\phi}{\partial x_n} + \lambda\frac{\partial f}{\partial x_n}\right)\Delta x_n \geq 0.$$

Substituting for Δx_n using (D.3.5) and combining the like terms, the above inequality can be rewritten as

$$\sum_{j=1}^{n-1}\left[\left(\frac{\partial \phi}{\partial x_j}+\lambda\frac{\partial f}{\partial x_j}\right)-\left(\frac{\partial f}{\partial x_n}\right)^{-1}\left(\frac{\partial f}{\partial x_j}\right)\left(\frac{\partial \phi}{\partial x_n}+\lambda\frac{\partial f}{\partial x_n}\right)\right]\Delta x_j \geq 0.$$

First choose λ such that

$$\frac{\partial \phi}{\partial x_n}+\lambda\frac{\partial f}{\partial x_n}=0. \tag{D.3.9}$$

This in turn requires that at the minimum

$$\sum_{j=1}^{n-1}\left(\frac{\partial \phi}{\partial x_j}+\lambda\frac{\partial f}{\partial x_j}\right)\Delta x_j \geq 0 \tag{D.3.10}$$

for all Δx_j, $1 \leq j \leq n-1$. Clearly, this can happen only if

$$\frac{\partial \phi}{\partial x_i}+\lambda\frac{\partial f}{\partial x_i}=0 \text{ for } 1 \leq i \leq n-1. \tag{D.3.11}$$

Combining (D.3.9) and (D.3.11) at the minimum, we have

$$\nabla\phi(\mathbf{x})+\lambda\nabla f(\mathbf{x})=\mathbf{0} \tag{D.3.12}$$

and hence

$$\delta\phi+\lambda\delta f=\mathbf{0}. \tag{D.3.13}$$

Rewriting (D.3.13) as

$$\delta(\phi+\lambda f)=0,$$

it can be verified that (D.3.12) is a condition for the unconstrained minimum of the new function

$$L(\mathbf{x},\lambda)=\phi(\mathbf{x})+\lambda f(\mathbf{x}) \tag{D.3.14}$$

called the **Lagrangian**, and λ is called the **Lagrangian multiplier**.

Thus the problem of minimizing a real-valued function of n variables under one constraint is converted to an unconstrained minimization of $L(\mathbf{x},\lambda)$ with $(n+1)$ variables. At the minimum, the derivatives of $L(\mathbf{x},\lambda)$ with respect to these $(n+1)$ variables must vanish which gives rise to

$$\nabla_{\mathbf{x}}L(\mathbf{x},\lambda)=\nabla\phi(\mathbf{x})+\lambda\nabla f(\mathbf{x})=0$$

which is (D.3.12) and

$$\nabla_{\lambda}L(\mathbf{x},\lambda)=f(\mathbf{x})=0$$

is the given constraint in (D.3.2). By solving these $(n+1)$ equations we can recover the values of $\mathbf{x}$ and λ corresponding to the minimum of $L(\mathbf{x},\lambda)$.

The following is an illustration of this method.

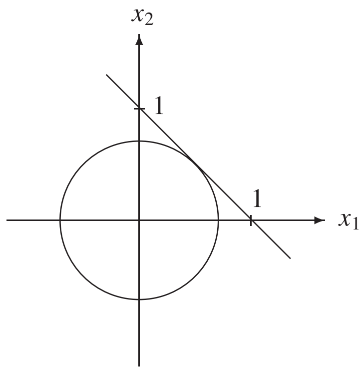

Fig. D.3.1 An illustration of Example D.3.1.

Example D.3.1(continued) Define

$$L(\mathbf{x}, \lambda) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1)$$
$$\nabla_x L(\mathbf{x}, \lambda) = \begin{bmatrix} 2x_1 + \lambda \\ 2x_2 + \lambda \end{bmatrix} = 0$$

from which we obtain $x_1 = x_2$ and $\lambda = -2x_2$. From

$$\nabla_\lambda L(\mathbf{x}, \lambda) = x_1 + x_2 - 1 = 0$$

we get $x_1 = x_2 = 1/2$ and $\lambda = -1$, at the minimum. The minimum value of $\phi(\mathbf{x}) = 1/2$. Refer to Figure D.3.1 for an illustration.

Special case A: Minimization of $\phi(\mathbf{x})$ along a line

Let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ and consider a line $\mathcal{L}$ in $\mathbb{R}^n$ passing through a point $\mathbf{y} \in \mathbb{R}^n$ in the direction $\mathbf{p} \in \mathbb{R}^n$. Any point $\mathbf{x}$ on $\mathcal{L}$ can be parametrically represented by

$$\mathbf{x} = \mathbf{y} + \alpha \mathbf{p}$$

for some $\alpha \in \mathbb{R}^n$. That is,

$$\mathbf{x} \in \mathcal{L} = \mathbf{y} + \text{Span}\{\mathbf{p}\}$$

where $\mathcal{L}$ is called the **affine** subspace of dimension one in $\mathbb{R}^n$. Our aim is to derive conditions for the minimum of $\phi(\mathbf{x})$ on $\mathcal{L}$. To this end, define $g : \mathbb{R} \longrightarrow \mathbb{R}$ as

$$g(\alpha) = \phi(\mathbf{y} + \alpha \mathbf{p}).$$

Then, the minimizer α_* of $g(\alpha)$ is obtained by solving

$$\frac{\mathrm{d}g}{\mathrm{d}\alpha} = \mathbf{p}^{\mathrm{T}} \nabla \phi(\mathbf{y} + \alpha \mathbf{p}) = 0.$$

If we denote $\mathbf{x}_* = \mathbf{y} + \alpha_* \mathbf{p}$, then the minimizer $\mathbf{x}_*$ of $\phi(\mathbf{x})$ is such that $\nabla \phi(\mathbf{x}_*)$ is orthogonal to the direction $\mathbf{p}$.

In the Example D.3.1, $\mathbf{p}$ is $(1,\ -1)^{\mathrm{T}}$ and $\mathbf{x}_* = (1/2, 1/2)^{\mathrm{T}}$, and $\nabla\phi(\mathbf{x}_*) = (1,\ 1)^{\mathrm{T}}$ and hence $\mathbf{p}^{\mathrm{T}}\nabla\phi(\mathbf{x}_*) = 0$.

This form of the conditions for optimality is used repeatedly in Chapters 9 through 11.

Lagrangian method: multiple equality constraints

We now generalize this Lagrangian method for minimizing $\phi(\mathbf{x})$ when Ω is specified by $k \geq 1$ constraint functions given in (D.3.1). Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)^{\mathrm{T}}$ be the vector of k Lagrangian multipliers. Construct the Lagrangian

$$L(\mathbf{x}, \lambda) = \phi(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}). \tag{D.3.15}$$

By extending the argument for the one constraint case, it can be verified that at the minimum $\nabla_{\mathbf{x}}\phi \equiv (\Delta\phi(x)/\Delta x, \partial\phi/\partial x_2, \ldots, \partial\phi/\partial x_n)^{\mathrm{T}}$, $\nabla_{\mathbf{x}} f \equiv (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n)^{\mathrm{T}}$, and $\nabla_{\mathbf{x}} L \equiv (\partial L/\partial x_1, \partial f/\partial x_2, \ldots,\ \partial L/\partial x_n)^{\mathrm{T}}$ satisfy

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) \equiv \nabla_{\mathbf{x}}\phi(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i \nabla_{\mathbf{x}} f_i(\mathbf{x}) = 0 \tag{D.3.16}$$

and

$$\nabla_{\lambda} L(\mathbf{x}, \lambda) = f_i(\mathbf{x}) = 0 \text{ for } 1 \leq i \leq k. \tag{D.3.17}$$

The values of the $(n + k)$ variables at the minimum are obtained as the solution of the $(n + k)$ equations in (D.3.16) and (D.3.17). Equation (D.3.16) implies that at the minimum, $\nabla_{\mathbf{x}}\phi(\mathbf{x})$ is a linear combination of the gradient $\nabla_{\mathbf{x}} f_i(\mathbf{x})$ of the constraint function $f_i(\mathbf{x})$ for $1 \leq i \leq k$.

Special case B: Minimization of ϕ in an affine subspace

Let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ and let $\{p_0, p_1, \ldots, p_{k-1}\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. For any $\mathbf{y} \in \mathbb{R}^n$ and for $\alpha_i \in \mathbb{R}$, $0 \leq i < k$, define

$$\begin{aligned} \mathcal{L}_k &= \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \mathbf{y} + \alpha_0 p_0 + \alpha_1 p_1 + \alpha_{k-1} p_{k-1}\} \\ &= \mathbf{y} + \mathrm{Span}\{p_0, p_1, \ldots, p_{k-1}\} \end{aligned}$$

called the affine subspace of dimension k in $\mathbb{R}^n$. Our aim is to minimize $\phi(\mathbf{x})$ on $\mathcal{L}_k$.

Let $\mathbf{P} \in \mathbb{R}^{n\times k}$, where

$$\mathbf{P} = \{p_0, p_1, \ldots, p_{k-1}\}$$

be the matrix built out of the k linearly independent column vectors defining $\mathcal{L}_k$. Let $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1})^{\mathrm{T}} \in \mathbb{R}^k$. Any $\mathbf{x} \in \mathcal{L}_k$ can be parametrically represented as

$$\mathbf{x} = \mathbf{y} + \mathbf{P}\alpha$$

for some vector $\alpha \in \mathbb{R}^k$. Define $G : \mathbb{R}^k \longrightarrow \mathbb{R}$ as

$$G(\alpha) = \phi(\mathbf{y} + \mathbf{P}\alpha).$$

Clearly, the minimizing vector α_* is obtained as the solution of

$$\mathbf{0} = \nabla G(\alpha) = \mathbf{P}^{\mathrm{T}} \nabla\phi(\mathbf{y} + \mathbf{P}\alpha)$$

$$= \begin{bmatrix} p_0^{\mathrm{T}} \\ p_1^{\mathrm{T}} \\ \vdots \\ p_{k-1}^{\mathrm{T}} \end{bmatrix} \nabla\phi(\mathbf{y} + \mathbf{P}\alpha).$$

If $\mathbf{x}_* = (\mathbf{y} + \mathbf{P}\alpha_*)$, then the gradient of $\phi(\mathbf{x})$ at $\mathbf{x}_*$, namely, $\nabla\phi(\mathbf{x}_*)$ is simultaneously orthogonal to each of the k directions $\{p_0, p_1, \ldots, p_{k-1}\}$, that is, $\nabla\phi(\mathbf{x}_*)$ is orthogonal to the affine subspace $\mathcal{L}_k$.

We conclude this discussion with a statement of a second-order sufficiency condition for a relative minimum of $\phi(\mathbf{x})$ under the constraints (D.3.1). Let $H(\mathbf{x}) = \nabla^2\phi(\mathbf{x})$ be the Hessian of $\phi(\mathbf{x})$. Likewise, let $F_i(\mathbf{x}) = \nabla^2 f_i(\mathbf{x})$ be the Hessian of $f_i(\mathbf{x})$ for $1 \le i \le k$, $G(\mathbf{x})$ is the corresponding Hessian of $L(\mathbf{x}, \lambda)$ in (D.3.15). Thus,

$$G(\mathbf{x}) = H(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i F_i(\mathbf{x}). \tag{D.3.18}$$

Define

$$T = \left\{\mathbf{y} | \mathbf{y}^{\mathrm{T}} \nabla f_i(\mathbf{x}) = 0, 1 \le i \le k\right\}.$$

That is, T consists of all vectors $\mathbf{y}$ that are orthogonal to $\nabla f_i(\mathbf{x})$, $1 \le i \le k$. T is indeed the plane that is tangent to $f_i(\mathbf{x})$, $1 \le i \le k$. Let $\mathbf{x}^*$ be such that there exists $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)^{\mathrm{T}}$, where

$$\begin{aligned} &\text{(a)} \nabla_{\mathbf{x}}\phi(\mathbf{x}^*) + \textstyle\sum_{i=1}^{k} \lambda_i \nabla_{\mathbf{x}} f_i(\mathbf{x}^*) = 0, \quad \textit{and} \\ &\text{(b) the matrix } G(\mathbf{x}^*) \text{ is positive definite on } T \end{aligned} \tag{D.3.19}$$

that is, for any $\mathbf{y} \neq \mathbf{0}$ and $\mathbf{y} \in T$, $\mathbf{y}G(\mathbf{x}^*)\mathbf{y} > 0$. Then, $\mathbf{x}^*$ is a relative minimum under the equality constraints.

The following is an illustration of the sufficiency condition.

Example D.3.2 Let $\mathbf{x} \in \mathbb{R}^2$, $\phi(\mathbf{x}) = x_1 + x_1x_2 + 3x_2^2$ and $f(\mathbf{x}) = x_1 + 2x_2 - 3 = 0$. The Lagrangian is given by $L(\mathbf{x}, \lambda) = \phi(\mathbf{x}) + \lambda f(\mathbf{x})$. The first-order necessary condition is given by

$$\begin{gathered} 1 + x_2 + \lambda = 0 \\ x_1 + 6x_2 + 2\lambda = 0, \quad \textit{and} \\ f(\mathbf{x}) = 0. \end{gathered}$$

Solving these equations, we get $x_1^* = 4$, $x_2^* = -1/2$ and $\lambda = -1/2$. It can be verified that

$$G(\mathbf{x}^*) = H(\mathbf{x}^*) = \begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix} \text{ and } F(\mathbf{x}^*) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since $\nabla f(\mathbf{x}^*) = (1, 2)^{\mathrm{T}}$, it can be verified that $T = \left\{ \frac{\alpha}{\sqrt{5}}(-2, 1)^{\mathrm{T}} \text{ for any real } \alpha \right\}$. From

$$\frac{\alpha}{\sqrt{5}}(-2, 1) \begin{pmatrix} 0 & 1 \\ 1 & 6 \end{pmatrix} \frac{\alpha}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \frac{2\alpha^2}{5} > 0$$

we conclude that $\mathbf{x}^* = (4, -1/2)$ is a strict relative minimum,

Remark D.3.1 It can be verified that the eigenvalues of

$$H(\mathbf{x}^*) = \begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix}$$

are given by $\mu_1 = \frac{6+\sqrt{40}}{2} > 0$ and $\mu_2 = \frac{6-\sqrt{40}}{2} < 0$, that is, $H(\mathbf{x}^*)$ is indefinite. In other words, one cannot characterize the nature of the extremum at $\mathbf{x}^*$ just by looking at $H(\mathbf{x}^*)$ alone. As shown above, $G(\mathbf{x}^*)$ in combination with tangent plane at $\mathbf{x}^*$ provides the clue for determining if $\mathbf{x}^*$ is a minimum.

D.4 Penalty function method

In this section we describe a second approach for converting the constrained minimization to an unconstrained problem using the **penalty function** method. The basic idea is to convert the original constrained problem into a sequence of unconstrained problems such that the minimizing solutions to these latter problems converge in the limit to the minimizing solution to the original problem.

Let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^{\mathrm{T}}$, where $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ for $1 \leq i \leq k$. Our aim is to minimize $\phi(\mathbf{x})$ subject to the constraints

$$f(\mathbf{x}) = \mathbf{0}. \tag{D.4.1}$$

Recall that $\mathcal{S} = \{\mathbf{x} | f(\mathbf{x}) = \mathbf{0}\} \subseteq \mathbb{R}^n$ is called the **feasible set**. In the Lagrangian formulation, the constraints are enforced at every step of the way. It turns out, however, algorithmically there is merit in seeking methods that violate feasibility so long as we can incorporate a penalty for each such violation – larger penalty for larger violation, where, for example, violation is measured by the distance of the vector $f(\mathbf{x})$ from the origin. Since the basic idea of minimization algorithm is strongly grounded on the **greedy principle** (refer to Chapter 9), in seeking the minimum, a cleverly designed algorithm would force the iterates in the direction of reducing the penalty which in turn force the iterates ever so closely towards the feasible set $\mathcal{S}$.

One useful way to quantify the penalty for violation of the constraints is to define a **quadratic** penalty

$$g(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{k} f_i^2(\mathbf{x}) = \frac{1}{2} f^{\mathrm{T}}(\mathbf{x}) f(\mathbf{x}), \tag{D.4.2}$$

which is the measure of the square of the Euclidean distance of the vector **f**(**x**) from the origin. Clearly, $g(\mathbf{x}) > 0$ when **x** is **not** feasible.

Let $\mathbb{R}^+ = \{x | x > 0\}$. Then, for $\rho \in \mathbb{R}^+$, define the penalty function $\psi : \mathbb{R}^n \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ as

$$\psi(\mathbf{x}, \rho) = \phi(\mathbf{x}) + \rho g(\mathbf{x}) = \phi(\mathbf{x}) + \frac{\rho}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x}). \tag{D.4.3}$$

The minimizer $\mathbf{x}_*(\rho)$ of $\psi(\mathbf{x}, \rho)$ is indeed the solution of

$$\begin{aligned} \mathbf{0} = \nabla_{\mathbf{x}}\psi(\mathbf{x}, \rho) &= \nabla\phi(\mathbf{x}) + \rho \textstyle\sum_{i=1}^{k} f_i(\mathbf{x}) \nabla f_i(\mathbf{x}) \\ &= \nabla\phi(\mathbf{x}) + \rho\mathbb{D}_f^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x}) \end{aligned} \tag{D.4.4}$$

where

$$\mathbb{D}_f(\mathbf{x}) = \left[\frac{\partial f_i}{\partial x_j}\right], 1 \le i \le k, 1 \le j \le n, \tag{D.4.5}$$

is the **Jacobian** (Appendix C) of **f**(**x**). The idea is to pick an increasing sequence $\{\rho_k\}$, that is

$$\rho_1 < \rho_2 < \rho_3 < \cdots < \rho_k < \cdots$$

such that

$$\lim_{k\to\infty} \rho_k = \infty \tag{D.4.6}$$

and solve for the minimum $\mathbf{x}_*(\rho_k)$ of the function $\psi(\mathbf{x}, \rho_k)$ by repeatedly solving (D.4.4). It can be shown under mild conditions (such as Condition **C** in Section D.1) that

$$\lim_{k\to\infty} \mathbf{x}_*(\rho_k) = \mathbf{x}_* \tag{D.4.7}$$

is the solution of the original constrained minimization problem.

Define a vector $\lambda(\rho) = (\lambda_1(\rho), \lambda_2(\rho), \ldots, \lambda_k(\rho))^{\mathrm{T}}$. We can rewrite (D.4.4) as

$$\begin{aligned} &\nabla\phi(\mathbf{x}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x})\lambda(\rho) \\ &\qquad = \nabla f_i(\mathbf{x}) + \textstyle\sum_{i=1}^{k} \lambda_i(\rho)\nabla f_i(\mathbf{x}) = \mathbf{0}. \end{aligned} \tag{D.4.8}$$

Comparing this with (D.3.16), it is tempting to conclude that $\lambda_i(\rho)$ plays a role similar to the Lagrangian multiplier. This similarity is more than “skin deep”. Indeed, it can be shown that

$$\lim_{k\to\infty} \lambda_*(\rho_k) = \rho_k f(\mathbf{x}_*(\rho_k)) = \lambda_*, \tag{D.4.9}$$

Given $\phi(\mathbf{x})$ to be minimized along with the constraint $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.
Let $\mathbf{x}_0 \in \mathbb{R}^n$ and $\rho_0 > 0$ be the initial choice.

For $k = 0, 1, 2, 3, \ldots$, do the following 3 steps:

Step 1 Solve (D.4.4) for the minimizer $\mathbf{x}_*(\rho_k)$ of $\psi(\mathbf{x}, \rho_k)$.

Step 2 If $\mathbf{f}(\mathbf{x}_*(\rho_k)) = \mathbf{0}$, then STOP.

Step 3 Define $\rho_{k+1} > \rho_k$, and go to Step 1.

Fig. D.4.1 Penalty function method.

is the Lagrangian multiplier that defines the constrained minimum $\mathbf{x}_*$ had we used the Lagrangian approach to begin with.

Example D.4.1 Let $\mathbf{x} = (x_1, x_2)^{\mathrm{T}}$ and $k = 1$. Consider minimizing $\phi(\mathbf{x}) = x_1^2 + x_2^2$ subject to $f(\mathbf{x}) = x_1 + x_2 - 1 = 0$. The penalty function is given by

$$\begin{aligned}\psi(\mathbf{x}, \rho) &= \phi(\mathbf{x}) + \tfrac{\rho}{2} f^2(\mathbf{x}) \\ &= (x_1^2 + x_2^2) + \tfrac{\rho}{2}(x_1 + x_2 - 1)^2.\end{aligned}$$

Solving

$$\mathbf{0} = \nabla\psi(\mathbf{x}, \rho) = \begin{bmatrix} x_1(2+\rho) + \rho x_2 - \rho \\ \rho x_1 + (2+\rho) - \rho \end{bmatrix}$$

we get

$$\mathbf{x}_*(\rho) = \left(\frac{1}{2+\rho^{-1}}, \frac{1}{2+\rho^{-1}}\right)^{\mathrm{T}}.$$

The multiplier

$$\lambda_*(\rho) = \rho f(\mathbf{x}_*(\rho)) = -\frac{1}{1+\rho^{-1}}.$$

By letting $\rho \longrightarrow \infty$, it follows that

$$\mathbf{x}_*(\rho) \longrightarrow \mathbf{x}_* = \left(\frac{1}{2}, \frac{1}{2}\right)^{\mathrm{T}}$$

and

$$\lambda_*(\rho) \longrightarrow \lambda_* = -1$$

which, from Example D.3.1 is the minimizer of the Lagrangian.

The algorithmic framework for the penalty function method is described in Figure D.4.1.

Remark D.4.1 The implementation of this idea is a bit tricky and challenging largely due to the requirement that for optimality, the penalty parameters grow without bound. To get a feel for this numerical instability, let us compute the condition number of the Hessian of the penalty function at the successive minima as the penalty parameter ρ_k increases with k. To simplify the discussion, let us consider the sample problem in Example D.4.1. It can be verified that the Hessian of $\psi(\mathbf{x}, \mathbf{p})$ at $(\mathbf{x}_*(\rho_k), \rho_k)$ is given by

$$\nabla^2\psi(\mathbf{x}_*(\rho_k), \rho_k) = \begin{bmatrix} (2+\rho_k) & \rho_k \\ \rho_k & (2+\rho_k) \end{bmatrix} \tag{D.4.10}$$

and its eigenvalues are $\lambda_1 = (2 + 2\rho_k)$ and $\lambda_2 = 2$, with $\lambda_1 > \lambda_2$ as $\rho_k > 0$. Hence, the condition number

$$\kappa(\rho) = \frac{\lambda_1}{\lambda_2} = 1 + \rho_k \tag{D.4.11}$$

which goes to infinity with ρ_k. However, while ρ_k grows unbounded, as $\mathbf{x}_*(\rho_k) \longrightarrow \mathbf{x}_*$, the true constrained minimum, the sequence $f(\mathbf{x}_*(\rho_k))$ move ever so close to the null vector such that their product $\lambda_*(\rho_k) = \rho_k f(\mathbf{x}_*(\rho_k))$ tends to a finite limit λ_*, the Lagrangian multiplier for the minimum $\mathbf{x}_*$.

Remark D.4.2 In the geophysical domain, the Lagrangian formulation has come to be known as the **strong constraint** formulation where feasibility is enforced all the time, and the penalty function approach is called the **weak constraint** formulation where feasibility is enforced only in the limit. The asymptotic result (D.4.9) provides a useful interpretation of the relation between these two useful ways of formulating and solving the constrained problems.

D.5 Augmented Lagrangian method

One way to reduce the impact of ill-conditioning in penalty methods is to consider a hybrid method that combines the penalty function with the classical Lagrangian method. In motivating this idea, let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^{\mathrm{T}}$, where $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, for $1 \le i \le k$. The problem is to

$$\text{Minimize } \phi(\mathbf{x}) \text{ when } \mathbf{f}(\mathbf{x}) = \mathbf{0}, \tag{D.5.1}$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)^{\mathrm{T}}$ is the Lagrangian multiplier. Since Lagrangian method enforces feasibility this problem is equivalent to

$$\text{Minimize } L(\mathbf{x}, \lambda) = \phi(\mathbf{x}) + \lambda^{\mathrm{T}}\mathbf{f}(\mathbf{x}) \text{ when } \mathbf{f}(\mathbf{x}) = \mathbf{0}. \tag{D.5.2}$$

We could now consider the penalty function approach for the latter problem. To this end, define the **augmented Lagrangian**

$$\eta : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^+ \longrightarrow \mathbb{R}$$

Given $\phi(\mathbf{x})$ to be minimized subject to the constraint $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.
Let $\mathbf{x}_0 \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}^k$ and $\rho_0 \in \mathbb{R}^+$.

For $k = 0, 1, 2, 3, \ldots$, do the following 4 steps:

Step 1 If $\nabla L(\mathbf{x}_k, \lambda_k) = \mathbf{0}$, then STOP.

Step 2 Let $\mathbf{x}_{k+1}$ be the minimizer of the unconstrained problem

$$\eta(\mathbf{x}, \lambda_k, \rho_k) = \phi(\mathbf{x}) + \lambda_k^{\mathrm{T}}\mathbf{f}(\mathbf{x}) + \frac{\rho_k}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x}_k)\mathbf{f}(\mathbf{x}_k)$$

Step 3 Define $\lambda_{k+1} = \lambda_k + \rho_k \mathbf{f}(\mathbf{x}_{k+1})$.

Step 4 Choose $\rho_{k+1} \geq \rho_k$, and go to Step 1.

Fig. D.5.1 Augmented Lagrangian Method.

as

$$\begin{aligned}\eta(\mathbf{x}, \lambda, \rho) &= L(\mathbf{x}, \lambda) + \tfrac{\rho}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x}) \\ &= \phi(\mathbf{x}) + \lambda^{\mathrm{T}}\mathbf{f}(\mathbf{x}) + \tfrac{\rho}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x})\end{aligned} \tag{D.5.3}$$

where ρ is the penalty parameter. Computing the gradient of η, we get

$$\begin{aligned}\nabla\eta(\mathbf{x}, \lambda, \rho) &= \nabla\phi(\mathbf{x}) + \textstyle\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}) + \rho \sum_{i=1}^{k} f_i(\mathbf{x}) \nabla f_i(\mathbf{x}) \\ &= \nabla\phi(\mathbf{x}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x})\lambda + \rho\mathbb{D}_f^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x}) \\ &= \nabla\phi(\mathbf{x}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x})(\lambda + \rho\mathbf{f}(\mathbf{x}).\end{aligned} \tag{D.5.4}$$

For a given ρ, the minimizing pair $(\mathbf{x}, \lambda)$ is obtained by solving

$$\nabla\phi(\mathbf{x}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x})(\lambda + \rho\phi(\mathbf{x})) = \mathbf{0}. \tag{D.5.5}$$

This development suggests the following framework for an algorithm described in Figure D.5.1.

From the definition, it follows that

$$\begin{aligned}\mathbf{0} &= \nabla_{\mathbf{x}}\eta(\mathbf{x}_{k+1}, \lambda_k, \rho_k) \\ &= \nabla\phi(\mathbf{x}_{k+1}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x}_{k+1})[\lambda_k \rho_k f(\mathbf{x}_{k+1})] \\ &= \nabla\phi(\mathbf{x}_{k+1}) + \mathbb{D}_f^{\mathrm{T}}(\mathbf{x}_{k+1})\lambda_{k+1} \\ &= \nabla_{\mathbf{x}} L(\mathbf{x}_{k+1}, \lambda_{k+1}).\end{aligned}$$

Hence, the algorithm will terminate when

$$\nabla_\lambda L(\mathbf{x}_{k+1}, \lambda_{k+1}) = \mathbf{f}(\mathbf{x}_{k+1}) = \mathbf{0}.$$

Example D.5.1 Let $\phi(\mathbf{x}) = x_1^2 + x_2^2$ and $f(\mathbf{x}) = x_1 + x_2 - 1$. Then,

$$\eta(\mathbf{x}, \lambda, \rho) = (x_1^2 + x_2^2) + \lambda(x_1 + x_2 - 1) + \frac{\rho}{2}(x_1 + x_2 - 1)^2$$

$$\nabla(\mathbf{x}, \lambda, \rho) = 2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \lambda\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \rho\begin{pmatrix} x_1 + x_2 - 1 \\ x_1 + x_2 - 1 \end{pmatrix}.$$

For $k = 0, 1, 2, 3, \ldots$, do the following:

Step 1 If $\nabla L(\mathbf{x}_k, \lambda_k) = \mathbf{0}$, then STOP.

Step 2 Compute $\mathbf{x}_{k+1} = (\mathbf{x}_{k+1,1}\mathbf{x}_{k+1,2})$ by solving the linear system

$$\begin{pmatrix} (2+\rho_k) & \rho_k \\ \rho_k & (2+\rho_k) \end{pmatrix} \begin{pmatrix} \mathbf{x}_{k+1,1} \\ \mathbf{x}_{k+1,2} \end{pmatrix} = \begin{pmatrix} \rho_k - \lambda_k \\ \rho_k - \lambda_k \end{pmatrix}$$

Step 3 Define $\lambda_{k+1} = \lambda_k + \rho_k \mathbf{f}(\mathbf{x}_{k+1})$.

Step 4 Choose $\rho_{k+1} \geq \rho_k$, and go to Step 1.

Fig. D.5.2

Setting $\nabla(\mathbf{x}, \lambda, \rho) = \mathbf{0}$, we obtain

$$\begin{pmatrix} (2+\rho) & \rho \\ \rho & (2+\rho) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \rho - \lambda \\ \rho - \lambda \end{pmatrix}.$$

We now state the algorithm. Pick $\mathbf{x}_0 = (x_{01}, x_{02})^{\mathrm{T}}$, λ_0, and $\rho_0 > 0$.

We encourage the reader to perform the iterations for various values of ρ and examine the effect of keeping ρ fixed.

Remark D.5.1 The primary advantage of the augmented Lagrangian approach is that the penalty factor ρ needs to be large enough and does **not** have to grow unbounded. This, in a sense, eliminates the instability resulting from the ill-conditioning of the Hessian matrix with increasing values of ρ (refer to Remark D.4.1). The question then is how large a value of ρ to use? It can be shown that any value ρ that is large enough to render the Hessian of

$$\psi(\mathbf{x}, \rho) = \phi(\mathbf{x}) + \frac{\rho}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x}) \tag{D.5.6}$$

positive definite on the domain of consideration is sufficient for the convergence of the algorithm given above. Indeed, the larger the value of ρ the better is the convergence rate but it will also increase the condition number of interest to us. Thus, there is a trade-off betwen avoiding numerical instability and speed of convergence.

The Hessian of the $\psi(\mathbf{x}, \rho)$ for the problem in Example D.5.1 is given by

$$\nabla^2\psi(\mathbf{x}, \rho) = \begin{bmatrix} (2+\rho) & \rho \\ \rho & (2+\rho) \end{bmatrix}.$$

The eigenvalues of this matrix are $\lambda_1 = 2 + \rho$ and $\lambda_2 = 2$ are positive (since $\rho > 0$). We readily see that this Hessian is positive definite. In other words, any positive value of ρ will work. We encourage the reader to examine the variation of speed of convergence as a function of ρ.

Notes and references

This appendix covers some of the basic results from the vast and ever growing corpus of literature dealing with optimization. For a derivation of conditions for optimality refer to Luenberger (1969) and (1973), Nash and Sofer (1996), and Hestenes (1975) and (1980). Our treatment of Lagrangian multipliers is adapted from Lanczos (1970). Dennis and Schnabel (1996) provide an elegant description of algorithms for unconstrained optimization. Penalty function methods are treated extensively in Nash and Sofer (1996). Augmented Lagrangian method was introduced independently by Hestenes and Powell in 1969. Refer to Nash and Sofer (1996), Pierre and Lowe (1975), and Hestenes (1975) and (1980) for detailed analysis of this class of methods.