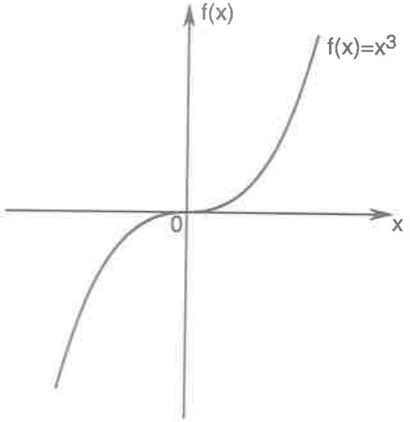

Figure 6.6 The point 0 satisfies the FONC and SONC, but is not a minimizer

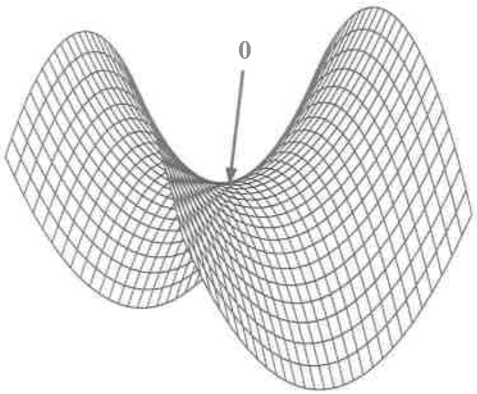

Figure 6.7 Graph of $f(\boldsymbol{x}) = x_1^2 - x_2^2$. The point $\mathbf{0}$ satisfies the FONC but not SONC; this point is not a minimizer

Theorem 6.3 ***Second-Order Sufficient Condition (SOSC), Interior Case.*** *Let $f \in \mathcal{C}^2$ be defined on a region in which $\boldsymbol{x}^*$ is an interior point. Suppose that*

1. $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$; *and*
2. $\boldsymbol{F}(\boldsymbol{x}^*) > 0$.

Then, $\boldsymbol{x}^*$ is a ~~strict~~ local minimizer of f. □

P34

Proof. Because $f \in \mathcal{C}^2$, we have $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{F}^T(\boldsymbol{x}^*)$. Using assumption 2 and Rayleigh's inequality it follows that if $\boldsymbol{d} \neq \mathbf{0}$, then $0 < \lambda_{\min}(\boldsymbol{F}(\boldsymbol{x}^*))\|\boldsymbol{d}\|^2 \leq \boldsymbol{d}^T\boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d}$. By Taylor's theorem and assumption 1,

$$f(\boldsymbol{x}^* + \boldsymbol{d}) - f(\boldsymbol{x}^*) = \frac{1}{2}\boldsymbol{d}^T\boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{d} + o(\|\boldsymbol{d}\|^2) \geq \frac{\lambda_{\min}(\boldsymbol{F}(\boldsymbol{x}^*))}{2}\|\boldsymbol{d}\|^2 + o(\|\boldsymbol{d}\|^2).$$

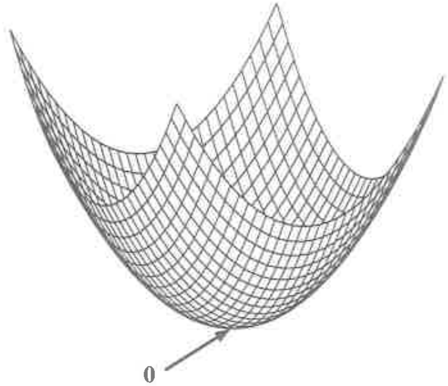

Figure 6.8 Graph of $f(\boldsymbol{x}) = x_1^2 + x_2^2$

Hence, for all $\boldsymbol{d}$ such that $\|\boldsymbol{d}\|$ is sufficiently small,

$$f(\boldsymbol{x}^* + \boldsymbol{d}) > f(\boldsymbol{x}^*),$$

and the proof is completed. ■

Example 6.7 Let $f(\boldsymbol{x}) = x_1^2 + x_2^2$. We have $\nabla f(\boldsymbol{x}) = [2x_1, 2x_2]^T = \mathbf{0}$ if and only if $\boldsymbol{x} = [0,0]^T$. For all $\boldsymbol{x} \in \mathbb{R}^2$, we have

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0.$$

The point $\boldsymbol{x} = [0,0]^T$ satisfies the FONC, SONC, and SOSC. It is a strict local minimizer. Actually $\boldsymbol{x} = [0,0]^T$ is a strict global minimizer. Figure 6.8 shows the graph of $f(\boldsymbol{x}) = x_1^2 + x_2^2$. ■

In this chapter, we presented a theoretical basis for the solution of nonlinear unconstrained problems. In the following chapters, we are concerned with iterative methods of solving such problems. Such methods are of great importance in practice. Indeed, suppose that one is confronted with a highly nonlinear function of 20 variables. Then, the FONC requires the solution of 20 nonlinear simultaneous equations for 20 variables. These equations, being nonlinear, will normally have multiple solutions. In addition, we would have to compute 210 second derivatives (provided $f \in \mathcal{C}^2$) to use the SONC or SOSC. We begin our discussion of iterative methods in the next chapter with search methods for functions of one variable.

6.1
6.5
6.7
6.14
6.18
6.22

EXERCISES

6.1 Consider the problem

$$\text{minimize} \quad f(\boldsymbol{x})$$

$$\text{subject to} \quad \boldsymbol{x} \in \Omega,$$

where $f \in \mathcal{C}^2$. For each of the following specifications for Ω, $\boldsymbol{x}^*$, and f, determine if the given point $\boldsymbol{x}^*$ is: (i) definitely a local minimizer; (ii) definitely not a local minimizer; or (iii) possibly a local minimizer. Fully justify your answer.

a. $f : \mathbb{R}^2 \to \mathbb{R}$, $\Omega = \{\boldsymbol{x} = [x_1, x_2]^T : x_1 \geq 1\}$, $\boldsymbol{x}^* = [1, 2]^T$, and gradient $\nabla f(\boldsymbol{x}^*) = [1, 1]^T$.

b. $f : \mathbb{R}^2 \to \mathbb{R}$, $\Omega = \{\boldsymbol{x} = [x_1, x_2]^T : x_1 \geq 1, x_2 \geq 2\}$, $\boldsymbol{x}^* = [1, 2]^T$, and gradient $\nabla f(\boldsymbol{x}^*) = [1, 0]^T$.

c. $f : \mathbb{R}^2 \to \mathbb{R}$, $\Omega = \{\boldsymbol{x} = [x_1, x_2]^T : x_1 \geq 0, x_2 \geq 0\}$, $\boldsymbol{x}^* = [1, 2]^T$, gradient $\nabla f(\boldsymbol{x}^*) = [0, 0]^T$, and Hessian $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{I}$ (identity matrix).

d. $f : \mathbb{R}^2 \to \mathbb{R}$, $\Omega = \{\boldsymbol{x} = [x_1, x_2]^T : x_1 \geq 1, x_2 \geq 2\}$, $\boldsymbol{x}^* = [1, 2]^T$, gradient $\nabla f(\boldsymbol{x}^*) = [1, 0]^T$, and Hessian

$$\boldsymbol{F}(\boldsymbol{x}^*) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

6.2 Show that if $\boldsymbol{x}^*$ is a global minimizer of f over Ω, and $\boldsymbol{x}^* \in \Omega' \subset \Omega$, then $\boldsymbol{x}^*$ is a global minimizer of f over Ω'.

6.3 Suppose that $\boldsymbol{x}^*$ is a local minimizer of f over Ω, and $\Omega \subset \Omega'$. Show that if $\boldsymbol{x}^*$ is an interior point of Ω, then $\boldsymbol{x}^*$ is a local minimizer of f over Ω'. Show that the same conclusion cannot be made if $\boldsymbol{x}^*$ is not an interior point of Ω.

6.4 Let $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x}_0 \in \mathbb{R}^n$, and $\Omega \subset \mathbb{R}^n$. Show that

$$\boldsymbol{x}_0 + \arg\min_{\boldsymbol{x} \in \Omega} f(\boldsymbol{x}) = \arg\min_{\boldsymbol{y} \in \Omega'} f(\boldsymbol{y}),$$

where $\Omega' = \{\boldsymbol{y} : \boldsymbol{y} - \boldsymbol{x}_0 \in \Omega\}$.

6.5 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given below:

$$f(\boldsymbol{x}) = \boldsymbol{x}^T \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} \boldsymbol{x} + \boldsymbol{x}^T \begin{bmatrix} 3 \\ 5 \end{bmatrix} + 6.$$

a. Find the gradient and Hessian of f at the point $[1, 1]^T$.

b. Find the directional derivative of f at $[1, 1]^T$ with respect to a unit vector in the direction of maximal rate of increase.

c. Find a point that satisfies the FONC (interior case) for f. Does this point satisfy the SONC (for a minimizer)?

6.6 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given below:

$$f(\boldsymbol{x}) = \boldsymbol{x}^T \begin{bmatrix} 2 & 5 \\ -1 & 1 \end{bmatrix} \boldsymbol{x} + \boldsymbol{x}^T \begin{bmatrix} 3 \\ 4 \end{bmatrix} + 7.$$

a. Find the directional derivative of f at $[0, 1]^T$ in the direction $[1, 0]^T$.

b. Find all points that satisfy the first-order necessary condition for f.

Does f have a minimizer? If it does, then find all minimizer(s); otherwise explain why it does not.

6.7 Consider the problem

$$\begin{array}{ll} \text{minimize} & -x_2^2 \\ \text{subject to} & |x_2| \leq x_1^2 \\ & x_1 \geq 0, \end{array}$$

where $x_1, x_2 \in \mathbb{R}$.

a. Does the point $[x_1, x_1]^T = \boldsymbol{0}$ satisfy the first-order necessary condition for a minimizer? That is, if f is the objective function, is it true that $\boldsymbol{d}^T \nabla f(\boldsymbol{0}) \geq 0$ for all feasible directions $\boldsymbol{d}$ at $\boldsymbol{0}$?

b. Is the point $[x_1, x_1]^T = \boldsymbol{0}$ a local minimizer, a strict local minimizer, a local maximizer, a strict local maximizer, or none of the above?

6.8 Consider the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(\boldsymbol{x}) = 5x_2$ with $\boldsymbol{x} = [x_1, x_2]^T$, and $\Omega = \{\boldsymbol{x} = [x_1, x_2]^T : x_1^2 + x_2 \geq 1\}$. Answer each of the following questions, showing complete justification.

a. Does the point $\boldsymbol{x}^* = [0, 1]^T$ satisfy the first-order necessary condition?

b. Does the point $\boldsymbol{x}^* = [0, 1]^T$ satisfy the second-order necessary condition?

c. Is the point $\boldsymbol{x}^* = [0, 1]^T$ a local minimizer?

6.9 Consider the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

where $\boldsymbol{x} = [x_1, x_2]^T$, $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(\boldsymbol{x}) = 4x_1^2 - x_2^2$, and $\Omega = \{\boldsymbol{x} : x_1^2 + 2x_1 - x_2 \geq 0, x_1 \geq 0, x_2 \geq 0\}$.

a. Does the point $\boldsymbol{x}^* = \mathbf{0} = [0, 0]^T$ satisfy the first-order necessary condition?

b. Does the point $\boldsymbol{x}^* = \mathbf{0}$ satisfy the second-order necessary condition?

c. Is the point $\boldsymbol{x}^* = \mathbf{0}$ a local minimizer of the given problem?

6.10 Suppose that we are given n real numbers, $x_1, \ldots, x_n$. Find the number $\bar{x} \in \mathbb{R}$ such that the sum of the squared difference between $\bar{x}$ and the above numbers is minimized (assuming the solution $\bar{x}$ exists).

6.11 An art collector stands at distance of x feet from the wall where a piece of art (picture) of height a feet is hung, b feet above his eyes, as shown in Figure 6.9.

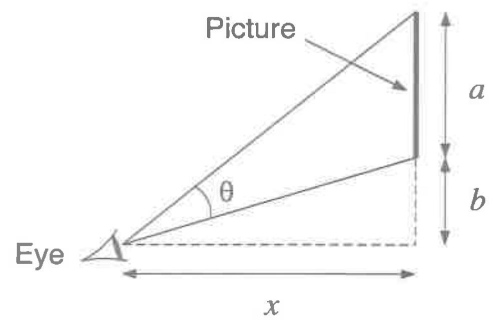

Figure 6.9 Art collector's eye position in Exercise 6.11

Find the distance from the wall for which the angle θ subtended by the eye to the picture is maximized.
Hint: (1) Maximizing θ is equivalent to maximizing $\tan(\theta)$;
(2) If $\theta = \theta_2 - \theta_1$, then $\tan(\theta) = (\tan(\theta_2) - \tan(\theta_1))/(1 + \tan(\theta_2)\tan(\theta_1))$.

6.12 Figure 6.10 shows a simplified model of a fetal heart monitoring system (the distances shown have been scaled down to make the calculations simpler). A heartbeat sensor is located at position x (see Figure 6.10).

The energy of the heartbeat signal measured by the sensor is the reciprocal of the squared distance from the source (baby's heart or mother's heart). Find the position of the sensor that maximizes the *signal-to-interference ratio*, which is the ratio of the signal energy from the baby's heart to the signal energy from the mother's heart.

6.13 An amphibian vehicle needs to travel from point A (on land) to point B (in water), as illustrated in Figure 6.11. The speeds at which the vehicle travels on land and water are v_1 and v_2, respectively.

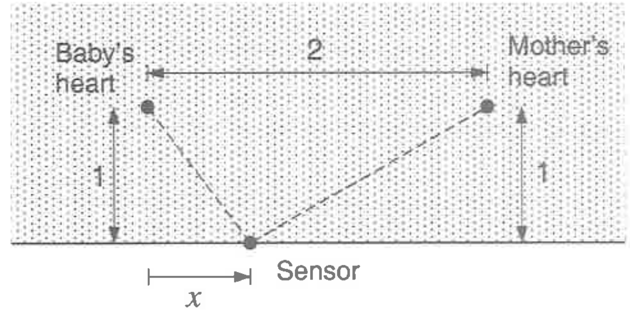

Figure 6.10 Simplified fetal heart monitoring system for Exercise 6.12

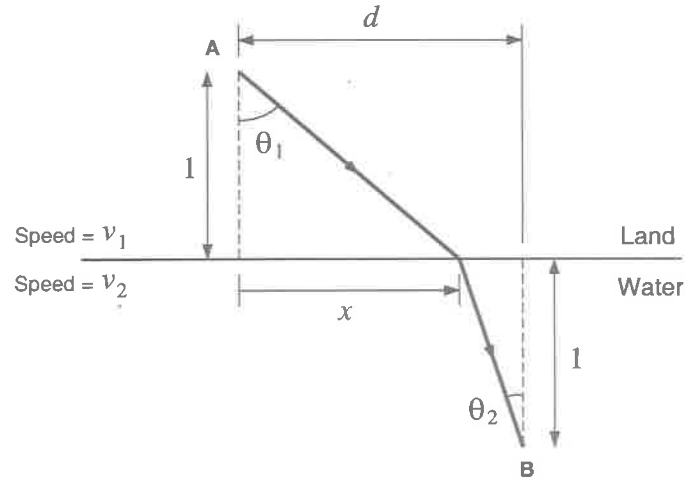

Figure 6.11 Path of amphibian vehicle in Exercise 6.13

a. Suppose that the vehicle traverses a path that minimizes the total time taken to travel from A to B. Use the first-order necessary condition to show that for the above optimal path, the angles θ_1 and θ_2 in the Figure 6.11 satisfy Snell's Law:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_1}{v_2}.$$

b. Does the minimizer for the problem in part a satisfies the second-order sufficient condition?

6.14 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(\boldsymbol{x}) = (x_1 - x_2)^4 + x_1^2 - x_2^2 - 2x_1 + 2x_2 + 1,$$

where $\boldsymbol{x} = [x_1, x_2]^T$. Suppose that we wish to minimize f over $\mathbb{R}^2$. Find all points satisfying the FONC. Do these points satisfy the SONC?

6.15 Show that if $\boldsymbol{d}$ is a feasible direction at a point $\boldsymbol{x} \in \Omega$, then for all $\beta > 0$, the vector $\beta\boldsymbol{d}$ is also a feasible direction at $\boldsymbol{x}$.

6.16 Let $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$. Show that $\boldsymbol{d} \in \mathbb{R}^n$ is a feasible direction at $\boldsymbol{x} \in \Omega$ if and only if $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$.

6.17 Let $f : \mathbb{R}^2 \to \mathbb{R}$. Consider the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & x_1, x_2 \geq 0, \end{array}$$

where $\boldsymbol{x} = [x_1, x_2]^T$. Suppose that $\nabla f(\boldsymbol{0}) \neq \boldsymbol{0}$, and

$$\frac{\partial f}{\partial x_1}(\boldsymbol{0}) \leq 0, \qquad \frac{\partial f}{\partial x_2}(\boldsymbol{0}) \leq 0.$$

Show that $\boldsymbol{0}$ cannot be a minimizer for the above problem.

6.18 Let $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{c} \neq \boldsymbol{0}$, and consider the problem of minimizing the function $f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x}$ over a constraint set $\Omega \subset \mathbb{R}^n$. Show that we cannot have a solution lying in the interior of Ω.

6.19 Consider the problem:

$$\begin{array}{ll} \text{maximize} & c_1x_1 + c_2x_2 \\ \text{subject to} & x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0, \end{array}$$

where c_1 and c_2 are constants such that $c_1 > c_2 \geq 0$. The above is a *linear programming* problem (see Part III). Assuming that the problem has an optimal feasible solution, use the *First-Order Necessary Conditions* to show that the *unique* optimal feasible solution $\boldsymbol{x}^*$ is $[1, 0]^T$.
Hint: First show that $\boldsymbol{x}^*$ cannot lie in the interior of the constraint set. Then, show that $\boldsymbol{x}^*$ cannot lie on the line segments $L_1 = \{\boldsymbol{x} : x_1 = 0, 0 \leq x_2 < 1\}$, $L_2 = \{\boldsymbol{x} : 0 \leq x_1 < 1, x_2 = 0\}$, $L_3 = \{\boldsymbol{x} : 0 \leq x_1 < 1, x_2 = 1 - x_1\}$.

6.20 *Line Fitting*. Let $[x_1, y_1]^T, \ldots, [x_n, y_n]^T$, $n \geq 2$, be points on the $\mathbb{R}^2$ plane (each $x_i, y_i \in \mathbb{R}$). We wish to find the straight line of "best fit" through these points ("best" in the sense that the average squared error is minimized); that is, we wish to find $a, b \in \mathbb{R}$ to minimize

$$f(a, b) = \frac{1}{n}\sum_{i=1}^{n}(ax_i + b - y_i)^2.$$

a. Let

$$\begin{aligned} \overline{X} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ \overline{Y} &= \frac{1}{n}\sum_{i=1}^{n} y_i \\ \overline{X^2} &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 \\ \overline{Y^2} &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 \\ \overline{XY} &= \frac{1}{n}\sum_{i=1}^{n} x_iy_i. \end{aligned}$$

Show that $f(a, b)$ can be written in the form $\boldsymbol{z}^T\boldsymbol{Q}\boldsymbol{z} - 2\boldsymbol{c}^T\boldsymbol{z} + d$, where $\boldsymbol{z} = [a, b]^T$, $\boldsymbol{Q} = \boldsymbol{Q}^T \in \mathbb{R}^{2\times 2}$, $\boldsymbol{c} \in \mathbb{R}^2$ and $d \in \mathbb{R}$, and find expressions for $\boldsymbol{Q}$, $\boldsymbol{c}$, and d in terms of $\overline{X}, \overline{Y}, \overline{X^2}, \overline{Y^2}$, and $\overline{XY}$.

b. Assume that the x_i, $i = 1, \ldots, n$, are not all equal. Find the parameters a^* and b^* for the line of best fit in terms of $\overline{X}, \overline{Y}, \overline{X^2}, \overline{Y^2}$, and $\overline{XY}$. Show that the point $[a^*, b^*]^T$ is the only local minimizer of f.
Hint: $\overline{X^2} - (\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2$.

c. Show that if a^* and b^* are the parameters of the line of best fit, then $\overline{Y} = a^*\overline{X} + b^*$ (and hence once we have computed a^*, we can compute b^* using the formula $b^* = \overline{Y} - a^*\overline{X}$).

6.21 Suppose that we are given a set of vectors $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(p)}\}$, $\boldsymbol{x}^{(i)} \in \mathbb{R}^n$, $i = 1, \ldots, p$. Find the vector $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ such that the average squared distance (norm) between $\bar{\boldsymbol{x}}$ and $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(p)}$,

$$\frac{1}{p}\sum_{i=1}^{p}\|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(i)}\|^2,$$

is minimized. Use the SOSC to prove that the vector $\bar{\boldsymbol{x}}$ found above is a strict local minimizer.

6.22 Consider a function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$ is a convex set and $f \in \mathcal{C}^1$. Given $\boldsymbol{x}^* \in \Omega$, suppose there exists $c > 0$ such that $\boldsymbol{d}^T\nabla f(\boldsymbol{x}^*) \geq c\|\boldsymbol{d}\|$ for all feasible directions $\boldsymbol{d}$ at $\boldsymbol{x}^*$. Show that $\boldsymbol{x}^*$ is a strict local minimizer of f over Ω.

6.23 Prove the following generalization of the second-order sufficient condition:
Theorem: Let Ω be a convex subset of $\mathbb{R}^n$, $f \in \mathcal{C}^2$ a real-valued function on Ω, and $\boldsymbol{x}^*$ a point in Ω. Suppose that there exists $c \in \mathbb{R}$, $c > 0$, such that for all feasible directions $\boldsymbol{d}$ at $\boldsymbol{x}^*$ ($\boldsymbol{d} \neq \boldsymbol{0}$), the following hold: