Practice Problems for nonlinear optimization

(Chong–Zak, 4th ed.)

6.2 Find minimizers and maximizers of the function

$$f(x_1, x_2) = \frac{1}{3}x_1^3 - 4x_1 + \frac{1}{3}x_2^3 - 16x_2.$$

6.8 Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}$:

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

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

b. Find the directional derivative of f at $[1, 1]^\top$ 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)?

(FONC: first order necessary condition: $\nabla f(x) = 0$

SONC: second order necessary condition: $D^2 f(x) \geq 0$)

6.10 Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}$:

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

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

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.

6.30 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 $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(p)}$,

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

Example 12.7 Find the point closest to the origin of $\mathbb{R}^3$ on the line of intersection of the two planes defined by the following two equations:

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1, \\ 4x_1 + x_2 + 3x_3 &= 0. \end{aligned}$$

Note that this problem is equivalent to the problem

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{x}\|^{2} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{array}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 & 3 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Thus, the solution to the problem is

$$\boldsymbol{x}^* = \boldsymbol{A}^\top (\boldsymbol{A}\boldsymbol{A}^\top)^{-1} \boldsymbol{b} = \begin{bmatrix} 0.0952 \\ 0.3333 \\ -0.2381 \end{bmatrix}.$$

12.19 Solve the problem

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{x} - \boldsymbol{x}_0\|^{2} \\ \text{subject to} & \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \boldsymbol{x} = 1, \end{array}$$

where $\boldsymbol{x}_0 = [0, -3, 0]^\top$.

An alternative method of arriving at the least-squares solution is to proceed as follows. First, we write

$$\begin{aligned} f(\boldsymbol{x}) &= \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 \\ &= (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^\top (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\ &= \frac{1}{2}\boldsymbol{x}^\top (2\boldsymbol{A}^\top \boldsymbol{A})\boldsymbol{x} - \boldsymbol{x}^\top (2\boldsymbol{A}^\top \boldsymbol{b}) + \boldsymbol{b}^\top \boldsymbol{b}. \end{aligned}$$

Therefore, f is a quadratic function. The quadratic term is positive definite because rank $\boldsymbol{A} = n$. Thus, the unique minimizer of f is obtained by solving the FONC (see Exercise 6.33); that is,

$$\nabla f(\boldsymbol{x}) = 2\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}^\top \boldsymbol{b} = \boldsymbol{0}.$$

The only solution to the equation $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$ is $\boldsymbol{x}^* = (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{b}$.

Example 12.1 Suppose that you are given two different types of concrete. The first type contains 30% cement, 40% gravel, and 30% sand (all percentages of weight). The second type contains 10% cement, 20% gravel, and 70% sand. How many pounds of each type of concrete should you mix together so that you get a concrete mixture that has as close as possible to a total of 5 pounds of cement, 3 pounds of gravel, and 4 pounds of sand?

The problem can be formulated as a least-squares problem with

$$\boldsymbol{A} = \begin{bmatrix} 0.3 & 0.1 \\ 0.4 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix},$$

where the decision variable is $\boldsymbol{x} = [x_1, x_2]^\top$ and x_1 and x_2 are the amounts of concrete of the first and second types, respectively. After some algebra, we obtain the solution:

$$\begin{aligned} \boldsymbol{x}^* &= (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{b} \\ &= \frac{1}{(0.34)(0.54) - (0.32)^2} \begin{bmatrix} 0.54 & -0.32 \\ -0.32 & 0.34 \end{bmatrix} \begin{bmatrix} 3.9 \\ 3.9 \end{bmatrix} \\ &= \begin{bmatrix} 10.6 \\ 0.961 \end{bmatrix}. \end{aligned}$$

(For a variation of this problem solved using a different method, see Example 15.7.) ∎

We now give an example in which least-squares analysis is used to fit measurements by a straight line.

Table 12.1 Experimental Data for Example 12.2.

i	0	1	2
t_i	2	3	4
y_i	3	4	15

Example 12.2 *Line Fitting.* Suppose that a process has a single input $t \in \mathbb{R}$ and a single output $y \in \mathbb{R}$. Suppose that we perform an experiment on the process, resulting in a number of measurements, as displayed in Table 12.1. The ith measurement results in the input labeled t_i and the output labeled y_i. We would like to find a straight line given by

$$y = mt + c$$

that fits the experimental data. In other words, we wish to find two numbers, m and c, such that $y_i = mt_i + c$, $i = 0, 1, 2$. However, it is apparent that there is no choice of m and c that results in the requirement above; that is, there is no straight line that passes through all three points simultaneously. Therefore, we would like to find the values of m and c that best fit the data. A graphical illustration of our problem is shown in Figure 12.2.

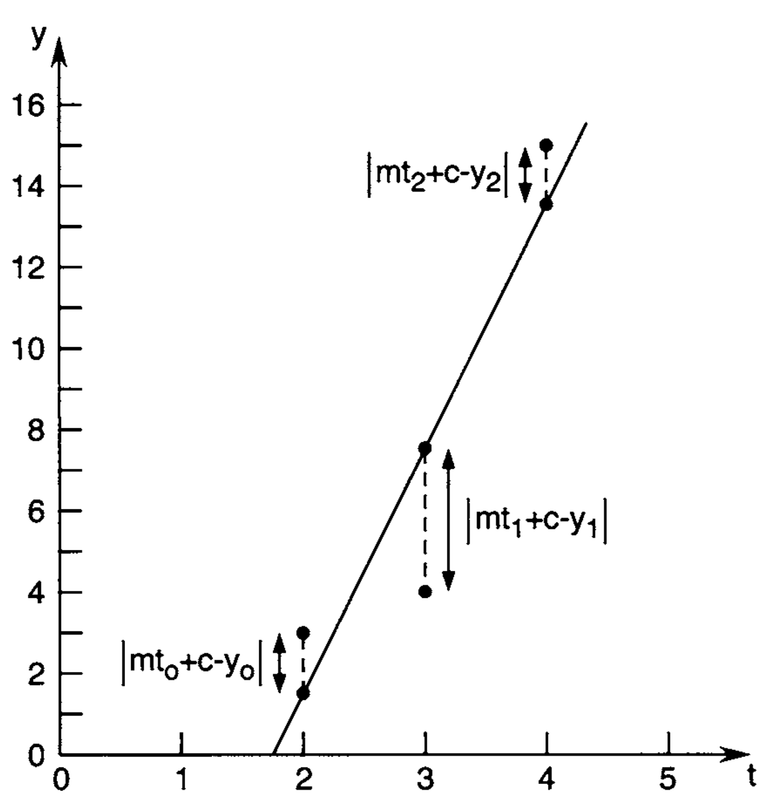

Figure 12.2 Fitting a straight line to experimental data.

We can represent our problem as a system of three linear equations of the form

$$\begin{aligned} 2m + c &= 3 \\ 3m + c &= 4 \\ 4m + c &= 15. \end{aligned}$$

We can write this system of equations as

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

where

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 3 \\ 4 \\ 15 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} m \\ c \end{bmatrix}.$$

Note that since

$$\operatorname{rank} \boldsymbol{A} < \operatorname{rank} [\boldsymbol{A}, \boldsymbol{b}],$$

the vector $\boldsymbol{b}$ does not belong to the range of $\boldsymbol{A}$. Thus, as we have noted before, the system of equations above is inconsistent.

The straight line of best fit is the one that minimizes

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 = \sum_{i=0}^{2} (mt_i + c - y_i)^2 .$$

Therefore, our problem lies in the class of least-squares problems. Note that the foregoing function of m and c is simply the total squared vertical distance (squared error) between the straight line defined by m and c and the experimental points. The solution to our least-squares problem is

$$\boldsymbol{x}^* = \begin{bmatrix} m^* \\ c^* \end{bmatrix} = (\boldsymbol{A}^\top \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{b} = \begin{bmatrix} 6 \\ -32/3 \end{bmatrix}.$$

Note that the error vector $\boldsymbol{e} = \boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{b}$ is orthogonal to each column of $\boldsymbol{A}$. ∎

Next, we give an example of the use of least-squares in wireless communications.

Example 12.3 *Attenuation Estimation.* A wireless transmitter sends a discrete-time signal $\{s_0, s_1, s_2\}$ (of duration 3) to a receiver, as shown in Figure 12.3. The real number s_i is the value of the signal at time i.

The transmitted signal takes two paths to the receiver: a direct path, with delay 10 and attenuation factor a_1, and an indirect (reflected) path, with delay 12 and attenuation factor a_2. The received signal is the sum of the signals from these two paths, with their respective delays and attenuation factors.

20.2 Find local extremizers for the following optimization problems:

a.
$$\begin{aligned} \text{Minimize} \quad & x_1^2 + 2x_1x_2 + 3x_2^2 + 4x_1 + 5x_2 + 6x_3 \\ \text{subject to} \quad & x_1 + 2x_2 = 3 \\ & 4x_1 + 5x_3 = 6. \end{aligned}$$

b.
$$\begin{aligned} \text{Maximize} \quad & 4x_1 + x_2^2 \\ \text{subject to} \quad & x_1^2 + x_2^2 = 9. \end{aligned}$$

c.
$$\begin{aligned} \text{Maximize} \quad & x_1x_2 \\ \text{subject to} \quad & x_1^2 + 4x_2^2 = 1. \end{aligned}$$

20.3 Find minimizers and maximizers of the function

$$f(\boldsymbol{x}) = (\boldsymbol{a}^\top \boldsymbol{x})(\boldsymbol{b}^\top \boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^3,$$

subject to

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_2 + x_3 &= 0, \end{aligned}$$

where

$$\boldsymbol{a} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

20.4 Consider the problem

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & h(\boldsymbol{x}) = 0, \end{aligned}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$, $h : \mathbb{R}^2 \to \mathbb{R}$, and $\nabla f(\boldsymbol{x}) = [x_1, x_1 + 4]^\top$. Suppose that $\boldsymbol{x}^*$ is an optimal solution and $\nabla h(\boldsymbol{x}^*) = [1, 4]^\top$. Find $\nabla f(\boldsymbol{x}^*)$.

20.5 Consider the problem

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{x} - \boldsymbol{x}_0\|^2 \\ \text{subject to} \quad & \|\boldsymbol{x}\|^2 = 9, \end{aligned}$$

where $\boldsymbol{x}_0 = [1, \sqrt{3}]^\top$.

a. Find all points satisfying the Lagrange condition for the problem.

20.7 Find local extremizers of

a. $f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 16$.

b. $f(x_1, x_2) = x_1^2 + x_2^2$ subject to $3x_1^2 + 4x_1x_2 + 6x_2^2 = 140$.

20.8 Consider the problem

$$\begin{array}{ll} \text{minimize} & 2x_1 + 3x_2 - 4, \qquad x_1, x_2 \in \mathbb{R} \\ \text{subject to} & x_1x_2 = 6. \end{array}$$

a. Use Lagrange's theorem to find all possible local minimizers and maximizers.

b. Use the second-order sufficient conditions to specify which points are strict local minimizers and which are strict local maximizers.

c. Are the points in part b global minimizers or maximizers? Explain.

20.9 Find all maximizers of the function

$$f(x_1, x_2) = \frac{18x_1^2 - 8x_1x_2 + 12x_2^2}{2x_1^2 + 2x_2^2}.$$

20.10 Find all solutions to the problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{x}^\top \begin{bmatrix} 3 & 4 \\ 0 & 3 \end{bmatrix} \boldsymbol{x} \\ \text{subject to} & \|\boldsymbol{x}\|^2 = 1. \end{array}$$