9

Introduction to Nonlinear Optimization

The class of nonlinear problems is very large and the models are formulated differently depending on how the nonlinear relations are expressed. There are many applications that require nonlinear models and some examples are discussed in Chapter 3. In this chapter (Section 9.1), we give some additional introductional examples.

In the case of nonlinear problems, unlike that of LP problems, there is no single method that can solve a general nonlinear model. The solution methods must be adopted to the special structure and to the type of functions used. Therefore, many different methods have been developed for various classes of nonlinear problems.

There are many alternatives for how to define different classes of nonlinear problems. One simple and often used method is to separate problems with and without constraints. The first class is called *unconstrained optimization* and the second *constrained optimization*. An alternative classification is to define problems with linear constraints and problems with nonlinear (or a mix of linear and nonlinear) constraints. If the problem has an quadratic objective function and linear constraints, we call it a *quadratic programming* problem. Solution methods for unconstrained and constrained optimization are described in Chapters 10 and 12, respectively.

Another general classification of nonlinear problems is between *convex problem* and *non-convex problem*. The theory and methods for checking convexity for functions, sets and problems is known as *convex analysis* and is treated in Section 9.3.

Most of the solution methods for nonlinear problems, and also the methods for investigating convexity, use some kind of approximation of functions. In Section 9.2 we therefore provide a description of the most used approximations.

9.1 Examples of nonlinear models

Least square problem

The company NZ Refining wants to establish a function that describes how the proportion of evaporated oil in the crude oil Kutubu2 depends on the temperature. This function is to be used in a larger production planning model. The company has measured the proportion of evaporated oil for a set of temperatures, and the result of this study is given in Table 9.1.

Measurement No.	Temperature (^oC)	Evaporated proportion (%)
1	20	0.0
2	30	5.8
3	80	14.7
4	125	31.6
5	175	43.2
6	225	58.3
7	275	78.4
8	325	89.4
9	360	96.4
10	420	99.1
11	495	99.5
12	540	99.9
13	630	100.0
14	700	100.0

Table 9.1 *Measurement data for the crude oil Kutubu2.*

The company believes that the behaviour can be described by a polynomial of the fourth degree. One side constraint is that the proportions of evaporated oil at the temperatures 20^o C and 700^o C must be 0% and 100%, respectively. Moreover, the proportion of evaporated oil can not decrease with higher temperature. Hence, we need an additional constraint which states that the derivative of the function must be positive. The problem is to determine the coefficients in the polynomial that best fit the measured data.

Formulation and model:
We introduce the following notations:

t_i = temperature at measurement point i, $i = 1, \ldots, 14$
F_i = proportion of evaporated oil at point i, $i = 1, \ldots, 14$

To establish a polynomial of order four we use the function

$$f(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$$

which approximates the proportion of evaporated oil as a function of temperature t. We can now formulate a nonlinear model where the coefficients a_0, a_1, a_2, a_3, and a_4 are unknown variables. The model can be stated as

$$\min z = \sum_{i=1}^{14} (F_i - (a_o + a_1 t_i + a_2 t_i^2 + a_3 t_i^3 + a_4 t_i^4))^2$$

s.t.

$$\begin{aligned} a_0 + 20a_1 + 20^2 a_2 + 20^3 a_3 + 20^4 a_4 &= 0 \quad (f(20) = 0) \\ a_0 + 700a_1 + 700^2 a_2 + 700^3 a_3 + 700^4 a_4 &= 100 \quad (f(700) = 100) \\ a_1 + 2a_2 t_i + 3a_3 t_i^2 + 4a_4 t_i^3 &\geq 0, \quad \forall\, i \ (f'(t) \geq 0) \end{aligned}$$

The first two constraints give the side constraints, and the third set of constraints states that the derivative of the approximative function should be non-negative. Apart from the constraints, we have a so called *least square problem*. The problem is to find coefficients in a polynomial such that the sum of the squared differences between the measured points F_i and the values of the polynomial in the points t_i is minimized.

Illustrations of two approximate functions are given in Figure 9.1. Here, each measured point is marked with an x. The solid line provides the best approximation found by solving the model above. The broken line provides the best approximation if the side constraints and the requirement on a positive derivative are removed. The latter represents the classical least square problem. Note that the figure is scaled with a factor of 100 with regard to both temperature and proportion of evaporated oil.

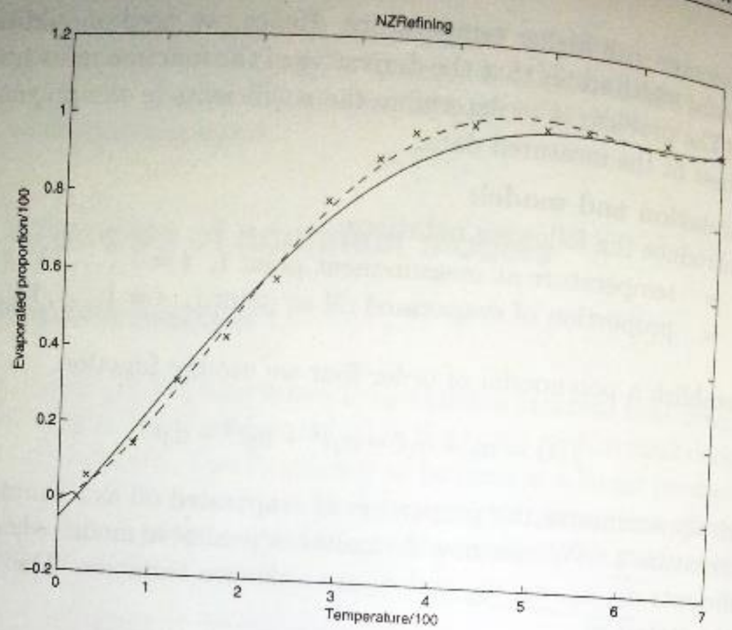

Figure 9.1 *Illustrations of two possible approximations of $f(t)$.*

Two dimensional cutting pattern

A company produces oil barrels in different dimensions. One part of the production process is to cut circular tops from rectangular pieces. The rectangular metal pieces available have the length L and width B. From each piece, the number of tops to cut is n. Top i has a given radius r_i. When the tops are cut, there is a requirement that there should be a distance d between them. The reason is to have some tolerance in the cutting. The same requirement of a tolerance d must also hold for each top in relation to any boundary.

The objective for the company is to place the tops as efficiently as possible. Unused parts of the metal piece at the edges with a rectangular form, can be used in another process and these parts have a value of p per area unit. Determine the placement of tops so that the remaining value of unused rectangular parts is maximized.

Figure 9.2 gives an example of a placement of five tops on a rectangular piece.

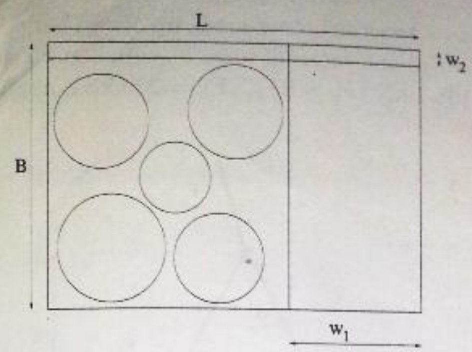

Figure 9.2 *Illustration of one cutting pattern.*

Formulation and model:

We define variables that describe the placement of each top's center, and the size of the unused rectangular piece. We introduce the following variables.

x_i = x-coordinate for the center of top i, $i = 1, \ldots, n$
y_i = y-coordinate for the center of top i, $i = 1, \ldots, n$
w_1 = length of the unused rectangular piece with the width B
w_2 = width of the unused rectangular piece with the length L

The nonlinear model can be stated as

$$\begin{aligned}
\max \quad & z = p(Bw_1 + Lw_2 - w_1 w_2) \\
\text{s.t.} \quad (x_i - x_j)^2 + (y_i - y_j)^2 &\geq (r_i + r_j + d)^2, \\
& \qquad i = 1, \ldots, n;\ j = 1, \ldots, n \ \ i \neq j \quad (1) \\
x_i + w_1 &\leq L - r_i - d, \qquad i = 1, \ldots, n \quad (2) \\
y_i + w_2 &\leq B - r_i - d, \qquad i = 1, \ldots, n \quad (3) \\
x_i &\geq r_i + d, \qquad i = 1, \ldots, n \quad (4) \\
y_i &\geq r_i + d, \qquad i = 1, \ldots, n \quad (5) \\
x_i,\ y_i &\geq 0, \qquad i = 1, \ldots, n \\
w_1,\ w_2 &\geq 0
\end{aligned}$$

Constraint (1) gives the pairwise distances between the centers of the tops. Constraints (2)-(5) define the smallest and largest values where the centers can be located. The problem is non-convex since the feasible region is non-convex. Very probably there are several local minima.

Construction of a bar-truss structure

Figure 9.3 shows a truss structure made up of two bars. The load case we study consists of one external force $\mathbf{F}$, with components $F_x = 24.8$ kN and $F_y = 198.4$ kN ($F_y = 8F_x, \|\mathbf{F}\| = 200$ kN), applied at the top of the structure.

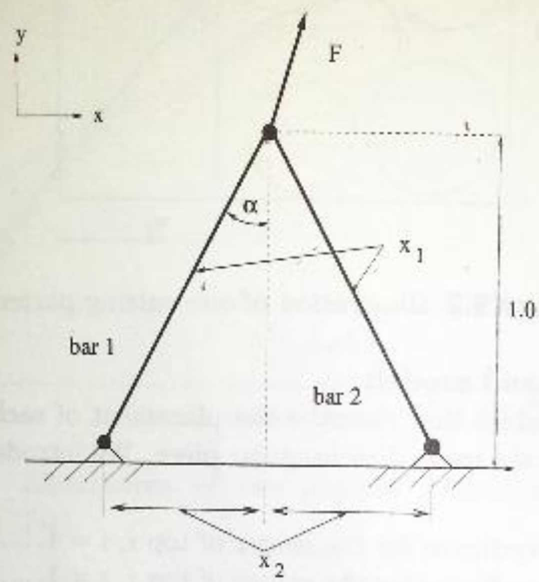

Figure 9.3 *Truss structure with two bars.*

We want to find the cross sectional areas of the bars and the distance between the support points so that the weight of the structure is minimized. There is a requirement that the stresses in the bars cannot exceed 100 N/mm^2 and that the bars have the same cross sectional areas. We have a lower and upper limit on the cross sectional areas of 0.2 cm^2 and 4.0 cm^2, respectively. Moreover, the distance between the supporting points must be at least 0.2 m and at most 3.2 m.

Formulation and model:
We define two variables as

x_1 = cross sectional area of the two bars (cm^2)
x_2 = half distance between the support points (m)

The length l of each bar is given by the relation $l^2 = x_2^2 + 1.0^2$. The forces in the bars denoted F_1 and F_2 are given by force equilibrium in the structure, i.e.

$$
\begin{aligned}
F_y - F_1 cos(\alpha) - F_2 cos(\alpha) &= 8F_x - (F_1 + F_2)\frac{1}{\sqrt{1+x_2^2}} &= 0 \\
F_x - F_1 sin(\alpha) + F_2 sin(\alpha) &= F_x - (F_1 - F_2)\frac{x_2}{\sqrt{1+x_2^2}} &= 0
\end{aligned}
$$

We can rewrite the equations and state the forces as

$$
\begin{aligned}
F_1 &= \left(\frac{\sqrt{1+x_2^2}}{2} + \frac{8\sqrt{1+x_2^2}}{2x_2}\right) F_x \\
F_2 &= \left(\frac{\sqrt{1+x_2^2}}{2} - \frac{8\sqrt{1+x_2^2}}{2x_2}\right) F_x
\end{aligned}
$$

The stresses in the bars are given by the quotient between force and area, i.e. $\sigma_1 = F_1/x_1$ and $\sigma_2 = F_2/x_1$. These expressions can be simplified to the expressions provided in the model below, where C is a parameter with value 12.4.

$$
\begin{array}{lrcll}
\min & z = w(x_1, x_2) = & 2x_1\sqrt{1+x_2^2} & & \\
\text{s.t.} & C\sqrt{1+x_2^2}(\frac{8}{x_1} + \frac{1}{x_1 x_2}) & \leq & 100 & \text{(bar 1)} \\
 & C\sqrt{1+x_2^2}(\frac{8}{x_1} - \frac{1}{x_1 x_2}) & \leq & 100 & \text{(bar 2)} \\
 & 0.2 \leq x_1 & \leq & 4.0 & \\
 & 0.1 \leq x_2 & \leq & 1.6 &
\end{array}
$$

The model is nonlinear and non-convex. One feasible solution is $x_1 = 1.41$ and $x_2 = 0.38$, with the stresses $\sigma_1 = 100$ in bar 1 and a total weight of the structure of $w = 3.02$. The constraint for stress in bar 2 is not active (stress $\sigma_2 < 100$).

9.2 Approximations of functions

To determine if a function or a set is convex, or to determine a search directions $\mathbf{d}^{(k)}$, we often use some approximation of the functions. In this section we describe how we can use first and second order approximations based on a Taylor series expansion. Other equivalent names for these approximations are linear and quadratic approximations. We assume that the functions are continuous and twice continuous differentiable.

We start with two familiar concepts from calculus. These are the gradient ∇f and the Hessian matrix $\mathbf{H}$ for a function $f(\mathbf{x})$. The Hessian matrix $\mathbf{H}$ is often denoted $\nabla^2 f$.

Definition 9.1

The gradient ∇f is a vector consisting of all partial derivatives to the function $f(\mathbf{x})$ and can be expressed as

$$\nabla f(x_1, \ldots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

□

Definition 9.2

The Hessian $\mathbf{H}$ is a quadratic matrix consisting of all partial second order derivatives to the function $f(\mathbf{x})$ and can be expressed as

$$\mathbf{H}(x_1, \ldots, x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

□

We use the notation f_1 for a first order approximation and f_2 for a second order approximation.

We start to study a one-dimensional function $f(x)$. A first order approximation of $f(x)$ around the point $\overline{x}$ is given by

$$f_1(x) = f(\overline{x}) + \frac{\partial f(\overline{x})}{\partial x}(x - \overline{x}) = f(\overline{x}) + f'(\overline{x})(x - \overline{x})$$

where $f'(\overline{x})$ is the derivative in the point $\overline{x}$. For an n-dimensional function $f(\mathbf{x})$ this can be generalized as follows.

Definition 9.3

A first order Taylor approximation of the function $f(\mathbf{x})$ around the point $\mathbf{x}^{(k)}$ is given by

$$f_1(\mathbf{x}) = f(\mathbf{x}^{(k)}) + \nabla f(\mathbf{x}^{(k)})^T(\mathbf{x} - \mathbf{x}^{(k)})$$

□

We can also express the approximation using component form as

$$f_1(\mathbf{x}) = f(\mathbf{x}^{(k)}) + \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_j}(x_j - x_j^{(k)})$$

Example 9.1

Determine a first order Taylor approximation of the function

$$f(x_1, x_2) = x_1^4 + x_1^2 + 2x_2^2 - 2x_1x_2$$

around the point $(x_1, x_2) = (1, 1)$.

Solution: *The gradient is* $\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1^3 + 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$

which in the point $(x_1, x_2) = (1, 1)$ *becomes* $\nabla f(1, 1) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

and we get

$$\begin{aligned} f_1(x_1, x_2) &= 2 + (4 \;\; 2)\begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} = 2 + 4(x_1 - 1) + 2(x_2 - 1) \\ &= 4x_1 + 2x_2 - 4 \end{aligned}$$

□

We can extend the first order approximation to a second order approximation by adding second order terms in the Taylor series expansion. We then get a quadratic approximation. We start with a one-dimensional function $f(x)$. A second order approximation of this function is given by

$$\begin{aligned} f_2(x) &= f(\bar{x}) + \frac{\partial f(\bar{x})}{\partial x}(x - \bar{x}) + \frac{1}{2}\frac{\partial^2 f(\bar{x})}{\partial x^2}(x - \bar{x})^2 \\ &= f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{1}{2}f''(\bar{x})(x - \bar{x})^2 \end{aligned}$$

where $f''(\bar{x})$ is the second derivative in the point $\bar{x}$. For an n-dimensional function $f(\mathbf{x})$ this can be generalized as follows.

Definition 9.4

A <u>*second order Taylor approximation*</u> *of the function* $f(\mathbf{x})$ *around the point* $\mathbf{x}^{(k)}$ *is given by*

$$\begin{aligned} f_2(\mathbf{x}) &= f(\mathbf{x}^{(k)}) + \nabla f(\mathbf{x}^{(k)})^T(\mathbf{x} - \mathbf{x}^{(k)}) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}^{(k)})^T \mathbf{H}(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) \end{aligned}$$

□

We can express this approximation using component form as

$$\begin{aligned} f_2(\mathbf{x}) &= f(\mathbf{x}^{(k)}) + \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_j}(x_j - x_j^{(k)}) \\ &+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_i \partial x_j}(x_i - x_i^{(k)})(x_j - x_j^{(k)}) \end{aligned}$$

Example 9.2

Determine a second order Taylor approximation to the function

$$f(x_1, x_2) = x_1^4 + x_1^2 + 2x_2^2 - 2x_1x_2$$

around the point $(x_1, x_2) = (1, 1)$. *(It is the same function as in Example 9.1.)*

Solution: *We have already computed the gradient* $\nabla f(x_1, x_2)$ *and the Hessian matrix is*

$$\mathbf{H}(x_1, x_2) = \begin{pmatrix} 12x_1^2 + 2 & -2 \\ -2 & 4 \end{pmatrix}$$

which in the point $(x_1, x_2) = (1, 1)$ *becomes* $\mathbf{H}(1, 1) = \begin{pmatrix} 14 & -2 \\ -2 & 4 \end{pmatrix}$ *and we get* $f_2(x_1, x_2) =$

$$2 + (4\ \ 2)\begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix}^T \begin{pmatrix} 14 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix}$$

$$= -4 + 4x_1 + 2x_2 + (7x_1^2 + 2x_2^2 - 2x_1x_2 - 12x_1 - 2x_2 + 7)$$

$$= 7x_1^2 + 2x_2^2 - 2x_1x_2 - 8x_1 + 3$$

□

9.3 Convex analysis

When we solve a nonlinear problem, it is very important to know if the problem is convex or not. The convexity characteristics decide which solution method is suitable to use and what solution quality we can expect when applying this method. If the problem is convex, each local optimum will also be a global optimum. Since most methods are search methods that only guarantee to find a local optimum, the convexity properties decide if we with certainty or not can announce that the solution we have found is the global optimum.

Definition 9.5
The problem

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \le b_i, \quad i = 1, \dots, m \end{array}$$

is a convex problem if $f(\mathbf{x})$ is a convex function and if the feasible region defined by the constraints $g_i(\mathbf{x}) \le b_i, i = 1, \dots, m$ is a convex set. □

The problem is convex also if we have a maximization problem and if the objective function is a concave function. Compared to the earlier definition of a convex problem (Definition 2.3 in Section 2.4), the feasible region X is in this definition defined using functions and inequalities, i.e. we have $X = \{\mathbf{x} \mid g_i(\mathbf{x}) \le b_i, \ i = 1, \dots, m\}$.

Theorem 9.1
If the problem

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in X \end{array}$$

is convex, then each local minimum is also a global minimum. □

Proof: Suppose the opposite, i.e. the point $\bar{\mathbf{x}}$ is a local minimum but there exists a point $\tilde{\mathbf{x}} \neq \bar{\mathbf{x}}, \ \tilde{\mathbf{x}} \in X$ where $f(\tilde{\mathbf{x}}) < f(\bar{\mathbf{x}})$. Since $f(\mathbf{x})$ is a convex function, we know that for an arbitrary point $\mathbf{x} = \lambda\bar{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{x}}$, where $0 \le \lambda \le 1$, we have

$$f(\lambda\bar{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{x}}) \le \lambda f(\bar{\mathbf{x}}) + (1-\lambda) f(\tilde{\mathbf{x}}) < \lambda f(\bar{\mathbf{x}}) + (1-\lambda) f(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$$

i.e. $f(\mathbf{x}) < f(\bar{\mathbf{x}})$. Let $\lambda \to 1$, i.e let $\mathbf{x} \to \bar{\mathbf{x}}$. All these points are feasible since X is a convex set. For $\mathbf{x}$ arbitrarily close to $\bar{\mathbf{x}}$ we still have $f(\mathbf{x}) < f(\bar{\mathbf{x}})$ which contradicts that $\bar{\mathbf{x}}$ is a local minimum. □

The problem in Definition 9.5 is a minimization problem with the feasible set expressed with $\le$-constraints only. All problems can be formulated in this form. The advantage is that we can then characterize the convexity of a problem through the properties of the functions defining the problem. The problem is convex if all functions $f(\mathbf{x})$ and $g_i(\mathbf{x}), i = 1, \dots, m$ are convex.

In Section 2.4 we defined a convex function and a convex set. However, it is difficult to determine convexity based on these definitions. Instead we use special methods to determine the convexity of a function or a set. First we show how convexity of a set can be determined by examining the functions defining the set. Then we present a set of theorems that can be used to determine whether or not a function is convex.

Convexity of a set

To determine if a set is convex we can use the following theorem.

Theorem 9.2
Let $X_1, X_2, \ldots, X_p$ be convex sets. Then, the intersection

$$X = X_1 \cap X_2 \cap \ldots \cap X_p$$

is a convex set.

□

The practical implication is that we can test convexity of a feasible region by testing if each individual constraint is convex. Theorem 9.2 is illustrated in Figure 9.4 for the case $p = 3$.

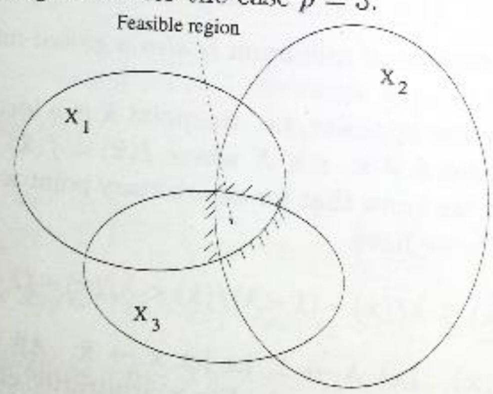

Figure 9.4 *The intersection of convex sets is a convex set.*

To show that a single constraint defines a convex set, we use the following theorem that is related to a single function.

Theorem 9.3
The set

$$X = \{\mathbf{x} \mid g(\mathbf{x}) \leq b\}$$

is a convex set if the function $g(\mathbf{x})$ is convex.

□

Proof: Let $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ be feasible points, i.e. $g(\mathbf{x}^{(1)}) \leq b$ and $g(\mathbf{x}^{(2)}) \leq b$. Select an arbitrary point $\mathbf{x}^{(3)} = \lambda\mathbf{x}^{(1)} + (1-\lambda)\mathbf{x}^{(2)}$ where $0 \leq \lambda \leq 1$. Since the function $g(\mathbf{x})$ is convex we have

$$g(\mathbf{x}^{(3)}) \leq \lambda g(\mathbf{x}^{(1)}) + (1-\lambda)g(\mathbf{x}^{(2)}) \leq \lambda b + (1-\lambda)b = b$$

which shows that $\mathbf{x}^{(3)}$ is a feasible point. Hence, the definition of a convex set is satisfied.

□

The conclusion is that we can use the methodology to test the convexity of a function also for testing convexity of sets.

If one or several constraints, independently, do not define a convex set, it is still possible that the intersection of the constraints (sets) defines a convex set. This is illustrated in Figure 9.5. However, it is very hard to prove that a set defined as the intersection of non-convex sets is a convex set. For some problems we can have the situation that it is impossible to comment on the convexity characteristics of the feasible set. In these cases, we cannot tell if the solution generated by the solution method is the global optimum solution or not.

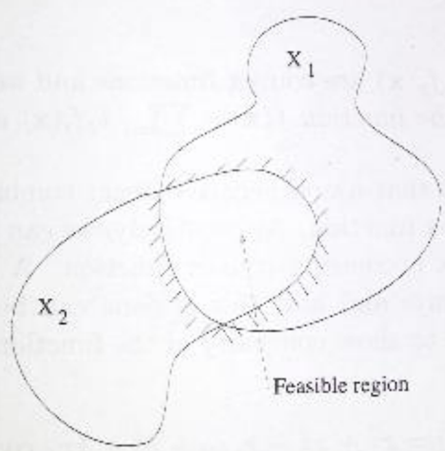

Figure 9.5 *Intersection of non-convex set can be convex.*

To show that a set is <u>not</u> convex, we can try to find an example (two feasible points) that simply contradicts the definition of a convex set.

Example 9.3
Is the set $X = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 2,\ x_1^2 - 4x_1 + x_2^2 \geq -2\}$ convex?

Solution: *Since the second set does not define a convex set, we may suspect that X is non-convex. Choose two points in X, e.g $\mathbf{x}^{(1)} =$*

$(1\ 1)^T$ *and* $\mathbf{x}^{(2)} = (1\ -1)^T$. *Let* $\mathbf{x}^{(3)} = \lambda\mathbf{x}^{(1)} + (1-\lambda)\mathbf{x}^{(2)}$, *where* $0 \le \lambda \le 1$. *Select, for example,* $\lambda = 0.5 \Rightarrow \mathbf{x}^{(3)} = (1\ 0)^T \notin X$. *Hence, the set* X *is not convex.* □

In practice, it may be difficult to find two points that can be used as a counter example. Often however, the only approach is to mathematically *show* that a set is not convex. Note the difference between showing that a set is not convex and not to be able to show that a set is convex. In the latter case, we still may have a convex set even if we failed to show it.

Convexity of a function

To determine if a function is convex, concave or neither convex nor concave, we must use special methods to examine the function. Again, we have little practical use of the formal definitions (Definition 2.4)

If the function consists of several parts, we can examine each part separately according to the following theorem.

Theorem 9.4
If $f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_p(\mathbf{x})$ *are convex functions and we have* $\lambda_i \ge 0$, $i = 1, \ldots, p$, *then the function* $f(\mathbf{x}) = \sum_{i=1}^{p} \lambda_i f_i(\mathbf{x})$ *is convex.* □

The theorem states that a non-negative linear combination of convex functions is a convex function. Alternatively, we can say that the sum of convex functions becomes a convex function. A function may be divided in many ways and how this is done can be decisive in how difficult it becomes to show convexity of the function.

Example 9.4
Is the function $f(\mathbf{x}) = x_1^2 + x_2^2 - x_1x_2 + x_3^4 + 4x_4$ *convex?*

Solution: $f(\mathbf{x})$ *can be expressed as the sum of three functions where* $f_1(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2$, $f_2(x_3) = x_3^4$ *and* $f_3(x_4) = 4x_4$. *Since all three functions are convex, the function* $f(\mathbf{x})$ *is also convex.* □

For each function, we can determine the convexity by studying the Hessian matrix. As support, we need the definition of a positive definite matrix.

Definition 9.6
The quadratic matrix $\mathbf{H}$ *is positive definite (positive semi-definite) if* $\mathbf{d}^T\mathbf{H}\mathbf{d} > 0$, *for all vectors* $\mathbf{d} \neq \mathbf{0}$ ($\mathbf{d}^T\mathbf{H}\mathbf{d} \geq 0, \ \forall\ \mathbf{d}$). □

We can now formulate the following useful theorem.

Theorem 9.5
Suppose the function $f(\mathbf{x})$ *is twice differentiable defined on a convex set* X. *Then we have:*

- $f(\mathbf{x})$ *is a convex function on* X *if the Hessian matrix* $\mathbf{H}$ *is positive semi-definite for all* $\mathbf{x} \in X$.
- $f(\mathbf{x})$ *is a strict convex function on* X *if the Hessian matrix* $\mathbf{H}$ *is positive definite for all* $\mathbf{x} \in X$.
- $f(\mathbf{x})$ *is a concave function on* X *if the Hessian matrix* $\mathbf{H}$ *is negative semi-definite for all* $\mathbf{x} \in X$.
- $f(\mathbf{x})$ *is a strict concave function on* X *if the Hessian matrix* $\mathbf{H}$ *is negative definite for all* $\mathbf{x} \in X$. □

This means we must be able to determine if the Hessian matrix $\mathbf{H}$ is positive semi-definite, positive definite, negative semi-definite or negative definite. If none of these hold, we have that $\mathbf{H}$ is *indefinite* and the function is neither convex nor concave.

One way to check the quadratic matrix $\mathbf{H}$ is to determine all eigenvalues to the matrix. These can be found, for example by finding the roots to the so called *characteristic equation* to $\mathbf{H}$, i.e. to the equation $\mathtt{det}(\mathbf{H} - \lambda I) = 0$. There are many efficient numerical methods for finding the eigenvalues to $\mathbf{H}$. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues to the Hessian matrix $\mathbf{H}$.

Theorem 9.6
If $\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$ *then* $\mathbf{H}$ *is positive semi-definite.*

If $\lambda_1 > 0, \ldots, \lambda_n > 0$ *then* $\mathbf{H}$ *is positive definite.*

If $\lambda_1 \leq 0, \ldots, \lambda_n \leq 0$ *then* $\mathbf{H}$ *is negative semi-definite.*

If $\lambda_1 < 0, \ldots, \lambda_n < 0$ *then* $\mathbf{H}$ *is negative definite.* □

There are also special methods for finding the smallest eigenvalue of a matrix. These are typically much faster than methods for finding all eigenvalues. If the smallest eigenvalue is strictly positive, then we know that all other eigenvalues are also strictly positive and that the matrix is positive definite.

Example 9.5
Is the function $f(x_1, x_2) = -x_1^2 - 4x_2^2 + x_1 x_2$ convex, concave or neither? The function and its level curves are given in Figure 9.6.

Solution: *The Hessian matrix to the function is* $\mathbf{H} = \begin{pmatrix} -2 & 1 \\ 1 & -8 \end{pmatrix}$ *and from the characteristic equation* $\det(\mathbf{H} - \lambda I) = 0$ *we get*

$$\begin{vmatrix} -2-\lambda & 1 \\ 1 & -8-\lambda \end{vmatrix} = 0 \text{ which gives } \lambda^2 + 10\lambda + 15 = 0$$

This equation has the roots $\lambda_1 = -5 + \sqrt{10}$ *and* $\lambda_2 = -5 - \sqrt{10}$. *Both roots are negative which implies that the Hessian matrix* $\mathbf{H}$ *is negative definite, and the function is a strictly concave function.* □

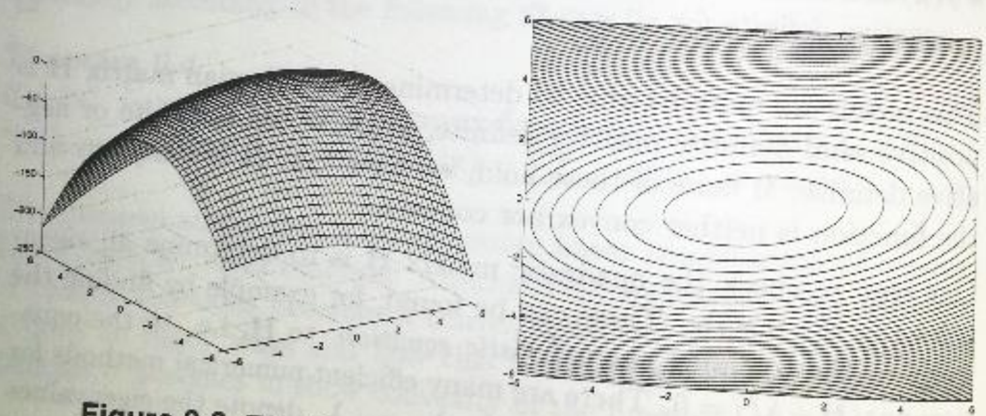

Figure 9.6 The function $f(x_1, x_2) = -x_1^2 - 4x_2^2 + x_1 x_2$.

An alternative approach to check if the matrix is positive or negative definite is to examine all *minor determinants* to $\mathbf{H}$. This approach can in practice be applied only to functions with few variables. It is a bit easier to examine if the function is positive definite or negative definite, i.e. if the function is strictly convex or strictly concave. In this case, it is enough to examine the n *leading minor determinants*, i.e. determinants of the quadratic submatrices

$$h_1 = \frac{\partial^2 f}{\partial x_1^2}, \quad \mathbf{h}_2 = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix},$$

$$\mathbf{h}_3 = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix}, \ldots\ldots, \mathbf{h}_n = \mathbf{H}$$

Theorem 9.7
$\mathbf{H}$ is positive definite if and only if

$$\det\ h_1 > 0,\ \det\ \mathbf{h}_2 > 0,\ \det\ \mathbf{h}_3 > 0,\ \ldots, \det\ \mathbf{H} > 0$$

$\mathbf{H}$ is negative definite if and only if

$$\det\ h_1 < 0,\ \det\ \mathbf{h}_2 > 0,\ \det\ \mathbf{h}_3 < 0, \ldots$$

i.e. if the sign of $\det\ \mathbf{h}_k$ *alternates between* < 0 *and* > 0 *(starting with* $\det\ h_1 < 0$*).* □

Example 9.6
Is the function $f(x_1, x_2) = x_1^4 + x_1^2 + 2x_2^2 - 2x_1x_2$ *convex, concave or neither? The function and its level curves are given in Figure 9.7.*

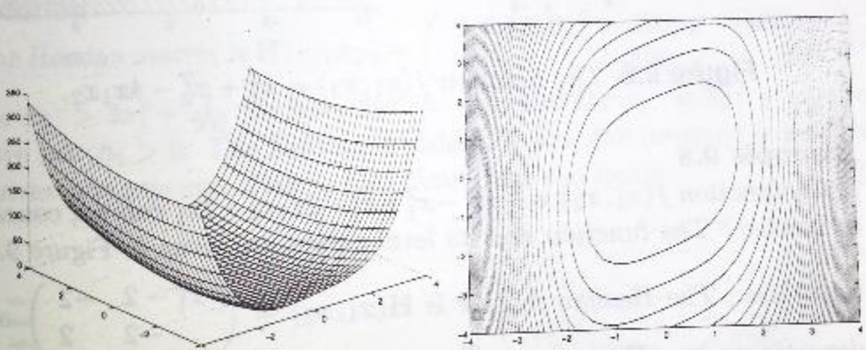

Figure 9.7 *The function* $f(x_1, x_2) = x_1^4 + x_1^2 + 2x_2^2 - 2x_1x_2$.

Solution*: The Hessian matrix is* $\mathbf{H}(x_1, x_2) = \begin{pmatrix} 12x_1^2 + 2 & -2 \\ -2 & 4 \end{pmatrix}$ *and we have* $\det\ h_1 = 12x_1^2 + 2 > 0$ *and* $\det\ \mathbf{H}(x_1, x_2) = 48x_1^2 + 4 > 0$ *for all values of* x_1 *and* x_2*. The conclusion is that* $\mathbf{H}$ *is positive definite and the function* $f(x_1, x_2)$ *is strictly convex.* □

It is easy to believe that $\mathbf{H}$ is semi-definite if we allow equalities in the relations in Theorem 9.7. However, this is not the case. Study for example the function $f(\mathbf{x}) = x_1^2 + 2x_2 - 3x_3^2$. The Hessian matrix has $\det h_1 = 2$, $\det \mathbf{h}_2 = 0$ and $\det \mathbf{h}_3 = 0$, but it is *not* positive semidefinite as the function is not convex.

Example 9.7

Is the function $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1x_2$ convex, concave or neither? The function and its level curves are given in Figure 9.8.

Solution: *The Hessian matrix is* $\mathbf{H}(x_1, x_2) = \begin{pmatrix} 2 & -4 \\ -4 & 2 \end{pmatrix}$, *and the eigenvalues are $\lambda_1 = 6$ and $\lambda_2 = -2$. The Hessian is indefinite and hence the function is neither convex nor concave.*

□

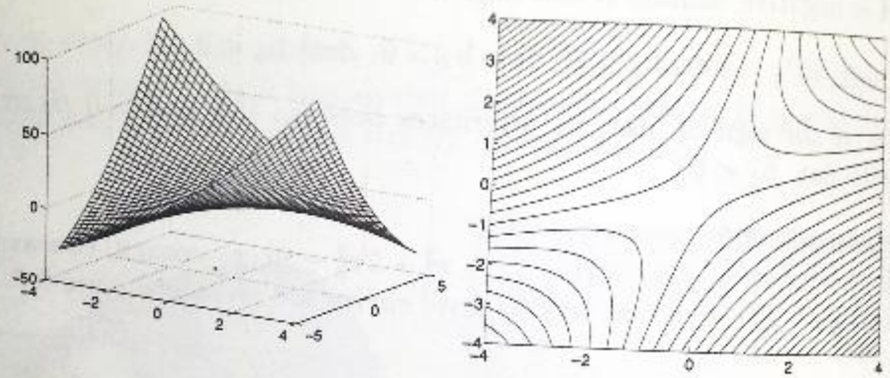

Figure 9.8 The function $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1x_2$.

Example 9.8

Is the function $f(x_1, x_2) = \frac{1}{12}x_1^4 - x_1^2 + x_2^2 - 2x_1x_2 - 2x_1$ convex, concave or neither? The function and its level curves are given in Figure 9.9.

Solution: The Hessian matrix is $\mathbf{H}(x_1, x_2) = \begin{pmatrix} x_1^2 - 2 & -2 \\ -2 & 2 \end{pmatrix}$ and depends on x_1. For $x_1 > \sqrt{2}$ or $x_1 < -\sqrt{2}$ we have $\det h_1 > 0$ and for $-\sqrt{2} < x_1 < \sqrt{2}$ we have $\det h_1 < 0$. The Hessian matrix is indefinite and the overall function is neither convex nor concave. Within the area $x_1 > 2$ or $x_1 < -2$, the function is convex.

□

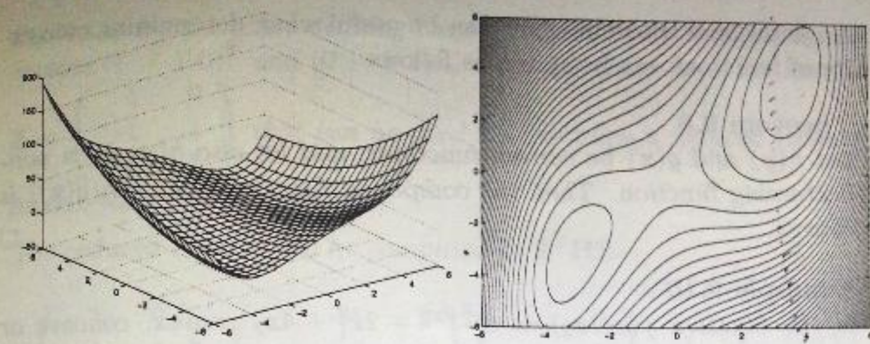

Figure 9.9 *The function* $f(x_1, x_2) = \frac{1}{12}x_1^4 - x_1^2 + x_2^2 - 2x_1x_2 - 2x_1$.

Example 9.9
Is the function $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ *convex, concave or neither? The function, which is called the Rosenbrock function, is often used to test the efficiency of various solution methods. The function is given in Figure 9.10.*

Solution*:*
The Hessian matrix is $\mathbf{H}(x_1, x_2) = \begin{pmatrix} 2 + 1200x_1^2 - 400x_2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$.
For $x_2 > 3x_1^2 + \frac{1}{200}$ *we have* det $h_1 < 0$ *and for* $x_2 < 3x_1^2 + \frac{1}{200}$ *we have* det $h_1 > 0$. *The Hessian is indefinite and the function is neither convex nor concave. This is also clear from the figure.* □

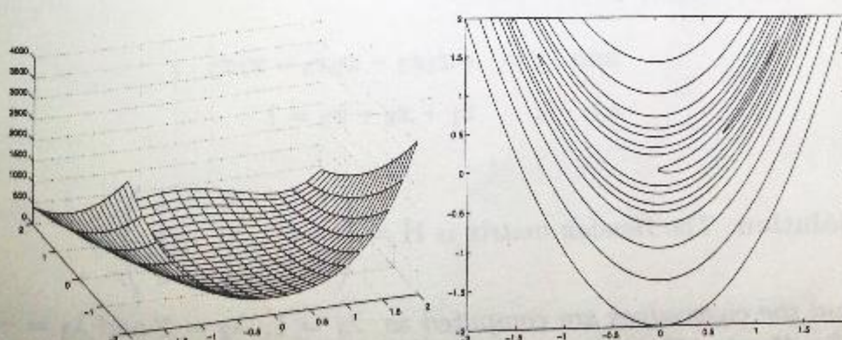

Figure 9.10 *The function* $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$.

An additional theorem which can be useful when determining convexity of functions can be stated as follows.

Theorem 9.8
Let $h(\mathbf{y})$ and $g(\mathbf{x})$ be convex functions, and let also $h(\mathbf{y})$ be a non-decreasing function. Then the compound function $f(\mathbf{x}) = h(g(\mathbf{x}))$ is convex.

□

Example 9.10
Is the function $f(x_1, x_2) = e^{x_1^2+x_2^2} + 2x_1^2 + 4x_2$ convex, concave or neither?

Solution: *The function is convex because*

1. $h(y) = e^y$ is a convex non-decreasing function and $g(x_1, x_2) = x_1^2 + x_2^2$ is a convex function, which according to Theorem 9.8 gives that the function $f_1(x_1, x_2) = e^{x_1^2+x_2^2}$ is convex.

2. The functions $f_2(x_1) = x_1^2$ and $f_3(x_2) = x_2$ are convex.

3. According to Theorem 9.4, we have a non-negative linear combination of convex functions.

□

For a minimization problem to be convex, it is enough to show that the function is convex over all points in the feasible region (see Theorem 9.5). The objective function does not need to be convex for *all* points in R^n. This is illustrated in the following example.

Example 9.11
Is the following problem convex?

$$\begin{array}{ll} \min & z = -x_1x_2 - x_2x_3 - x_1x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1 \end{array}$$

Solution: *The Hessian matrix is* $\mathbf{H} = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$

and the eigenvalues are computed as $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = -2$. The Hessian $\mathbf{H}$ is indefinite and the function is neither convex nor concave. However, we are only interested of the function for points in the feasible region and we therefore want to study the Hessian on the set $X = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 1\}$.

One basis for the plane defining the feasible solutions is given by the vectors $(1\ -1\ 0)^T$ *and* $(0\ 1\ -1)^T$. *Together these form the matrix* $\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}$ *that can be used to determine a reduced Hessian matrix which we can examine.*

The reduced Hessian can be computed as $\mathbf{Z}^T\mathbf{H}\mathbf{Z} =$

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}^T \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

The reduced Hessian matrix has eigenvalues $\lambda_1 = 1$ *and* $\lambda_2 = 3$ *and is therefore positive definite. This means that the function is convex over the feasible region, and hence the problem is convex.* □