Introduction to Nonlinear Optimization; and Optimality Conditions for Unconstrained Optimization Problems

Robert M. Freund

March, 2014

©2014 Massachusetts Institute of Technology. All rights reserved.
1 Introduction to Nonlinear Optimization

1.1 What is Nonlinear Optimization?

A nonlinear optimization problem is a problem of the following general form:

\[
\begin{align*}
(P:) & \quad \text{minimize}_x \quad f(x) \\
\text{s.t.} \quad & \quad x \in \mathcal{F}, \\
& \quad x \in \mathbb{R}^n.
\end{align*}
\]

Here \(x = (x_1, \ldots, x_n)^T\) are the variables, or unknowns, and are usually referred to as the decision variables. The function \(f(\cdot) : \mathbb{R}^n \to \mathbb{R}\) is called the objective function and it is this function that we seek to minimize, or more generally, to optimize. The expression \(x \in \mathcal{F}\) states that the decision variables \(x\) must lie in some specified set \(\mathcal{F}\). If the goal is maximization instead of minimization, then this can be formulated by either replacing “minimize” by “maximize” above, or by simply replacing \(f(\cdot)\) by \(-f(\cdot)\).

1.2 History of Nonlinear Optimization, in Brief

Most of the theoretical and computational interest in nonlinear optimization has taken place since 1947. However, it is useful to note that nonlinear optimization was first studied as early as the 1600s. Indeed, both Fermat (1638) and Newton (1670) studied the 1-dimensional nonlinear optimization problem:

\[
\min_x f(x) \quad \text{where } x \in \mathbb{R},
\]

and Newton developed the classical optimality condition: \(\frac{df(x)}{dx} = 0\). This was generalized by Euler (1755) to multivariate nonlinear optimization:

\[
\min f(x_1, \ldots, x_n)
\]

with the optimality condition: \(\nabla f(x) = 0\).

In the late 1700s both Euler and Lagrange examined problems in infinite dimensions and developed the calculus of variations. The study of optimization received very little attention thereafter until 1947, when modern linear
optimization was developed by Dantzig. In the late 1940s and early 1950s Kuhn and Tucker (preceded by Karush) developed general optimality conditions for nonlinear constrained optimization. At that time, applications of nonlinear optimization included problems in chemical engineering, portfolio optimization, and a host of decision problems in management and industrial operations. The 1960s saw the development of larger scale nonlinear optimization coincident with the expanding capabilities of computers. In 1984, new possibilities in optimization methods were developed with Karmarkar’s interior-point method for linear optimization, which was extended to nonlinear convex optimization by Nesterov and Nemirovskii in 1994. In the 1990s the important fields of semidefinite optimization, conic optimization, and “robust optimization” were developed. Since 2000 there has been great interest and progress in “huge-scale” nonlinear optimization, which considers problems in thousands or even millions of decision variables and constraints.

1.3 Some Applications of Optimization

Applications of nonlinear optimization arise in just about every human endeavor. (In fact, one could argue that as individuals and/or societies we constantly seek to make the choices that will optimize our own utility or happiness, or to optimize what is best for our society.) Nonlinear optimization problems arise throughout our industrial economy. For example, optimization problems arise in urban vehicle traffic management, in airline transportation management, and in the logistics of determining how to get goods and/or people to their destinations at least cost. In airline transportation, optimization arises in everything from determining the shape of an aircraft wing (and its skeleton) to optimize its lift/weight, to determining which crew should be assigned to which flight routes to minimize costs. Optimization is used to determine prices and inventories of products in stores and in online enterprises, as well as to compute advertising strategies and online banner advertisements for websites. In the finance industry, nonlinear optimization is used to compute efficient asset portfolios, to price financial instruments, and to determine trade strategies of investment funds. In data mining, optimization is used to efficiently decompose data into its principle components, to classify data by properties, and to predict everything from economic trends to the most desirable next purchase of an individual
consumer at an online retailer (recommender systems). In engineering, nonlinear optimization is used to compute antenna designs, sensor networks, optimal control policies, etc. In metamaterial design optimization is used to design materials with specific wave management properties. In physical systems, optimization is a governing principle behind physical laws – indeed many such laws arise as the solution of minimum energy principles. The list of applications of nonlinear optimization is virtually endless. Let us now visit several important applications that illustrate the structure, usefulness, and scope of nonlinear optimization.

1.3.1 Pricing Models

Consider a demand function \( d(p) \) for a single good as a function of the price \( p \) of the good, whose unit cost of production is a scalar \( c \). We presume that \( d(p) \) is a decreasing function, and we model it, perhaps locally, as:

\[
d(p) = \bar{d} - \bar{b}p,
\]

where \( \bar{b} > 0 \). If \( C \) is our production capacity, we seek to maximize contribution to earnings by solving:

\[
(P) \quad \max_p \quad f(p) := (\bar{d} - \bar{b}p) \times (p - c)
\]

s.t.

\[
\bar{d} - \bar{b}p \geq 0
\]

\[
p \geq 0
\]

\[
\bar{d} - \bar{b}p \leq C.
\]

Now let us generalize this model to \( n \) differentiated substitutable products. Demand for our goods is a function of the vector of prices \( p = (p_1, \ldots, p_n) \), and we presume that \( d(p) := d_1(p), \ldots, d_n(p) \) are linear functions of the
vector \( p \) of prices:

\[
d_1(p) = \bar{d}_1 - \bar{b}_{11}p_1 - \bar{b}_{12}p_2 - \ldots - \bar{b}_{1n}p_n \\
\vdots \\
d_n(p) = \bar{d}_n - \bar{b}_{n1}p_1 - \bar{b}_{n2}p_2 - \ldots - \bar{b}_{nn}p_n.
\]

We write this in vector form as:

\[
d(p) = \bar{d} - \bar{B}p.
\]

Let the per unit cost of production of good \( j \) be \( c_j \) and let \( C_j \) denote the production capacity of good \( j \) for \( j = 1, \ldots, n \). We write the vector of costs as \( c = (c_1, \ldots, c_n) \) and the vector of production capacities as \( C = (C_1, \ldots, C_n) \). We seek to maximize contribution to earnings by solving:

\[
(P) \quad \max_p \quad f(p) := (\bar{d} - \bar{B}p)^T (p - c)
\]

s.t. \( \bar{d} - \bar{B}p \geq 0 \)

\( p \geq 0 \)

\( \bar{d} - \bar{B}p \leq C. \)

1.3.2 A Pattern Classification Training Problem

By way of illustration, suppose we are given a set of pixilated digital images of faces. Suppose we have \( k \) male face images, expressed in digital form as \( a^1, \ldots, a^k \in \mathbb{R}^n \) (where each of the \( n \) coefficients corresponds to a pixel), and \( m \) female face images expressed in digital form as \( b^1, \ldots, b^m \in \mathbb{R}^n \). We would like to compute a linear decision rule comprised of a vector \( v \in \mathbb{R}^n \) and a scalar \( \beta \) that will satisfy \( v^T a^i > \beta \) for all \( i = 1, \ldots, k \) and \( v^T b^i < \beta \) for all \( i = 1, \ldots, m \). If we can compute \( v, \beta \) that satisfy these conditions, we can then use \( v, \beta \) to establish a prediction procedure to declare the gender of any other digital face image as follows: if we are given a face image \( c \in \mathbb{R}^n \), we will declare whether \( c \) is the image of a male or female face as follows:
• if $v^T c > \beta$, then we declare that $c$ is a male face image, and
• if $v^T c < \beta$, then we declare that $c$ is a female face image.

Much more generally, suppose we are given:

• points $a^1, \ldots, a^k \in \mathbb{R}^n$ that have property “P”, and
• points $b^1, \ldots, b^m \in \mathbb{R}^n$ that do not have property “P”.

We would like to use these $k + m$ points to develop a linear rule that can be used to predict whether or not other points $c$ might or might not have property $P$. In particular, we seek a vector $v$ and a scalar $\beta$ for which:

• $v^T a^i > \beta$ for all $i = 1, \ldots, k$
• $v^T b^i < \beta$ for all $i = 1, \ldots, m$.

We will then use $v, \beta$ to predict whether or not any other point $c$ has property $P$ or not. If we are given another vector $c$, we will declare whether $c$ has property $P$ or not as follows:

• if $v^T c > \beta$, then we declare that $c$ has property $P$, and
• if $v^T c < \beta$, then we declare that $c$ does not have property $P$.

We therefore seek $v, \beta$ that define the hyperplane

$$H_{v, \beta} := \{ x : v^T x = \beta \}$$

for which:

• $v^T a^i > \beta$ for all $i = 1, \ldots, k$
• $v^T b^i < \beta$ for all $i = 1, \ldots, m$. 
This is illustrated in Figure 1.

In addition, we would like the hyperplane $H_{v,\beta}$ to be as far away from the points $a^1, \ldots, a^k, b^1, \ldots, b^k$ as possible. Elementary analysis shows that the distance of the hyperplane $H_{v,\beta}$ to any point $a^i$ is equal to

$$\frac{v^T a^i - \beta}{\|v\|}$$

and similarly the distance of the hyperplane $H_{v,\beta}$ to any point $b^i$ is equal to

$$\frac{\beta - v^T b^i}{\|v\|}.$$

If we normalize the vector $v$ so that $\|v\| = 1$, then the minimum distance of the hyperplane $H_{v,\beta}$ to the points $a^1, \ldots, a^k, b^1, \ldots, b^k$ is then:

$$\min\{v^T a^1 - \beta, \ldots, v^T a^k - \beta, \beta - v^T b^1, \ldots, \beta - v^T b^m\}.$$  

Therefore we would also like $v$ and $\beta$ to satisfy:

- $\|v\| = 1$, and
- $\min\{v^T a^1 - \beta, \ldots, v^T a^k - \beta, \beta - v^T b^1, \ldots, \beta - v^T b^m\}$ is maximized.

This is formulated as the following optimization problem:
\[PCP: \ \text{maximize}_{v, \beta, \delta} \quad \delta\]
\[
\text{s.t.} \quad v^T a^i - \beta \quad \geq \quad \delta, \quad i = 1, \ldots, k
\]
\[
\beta - v^T b^i \quad \geq \quad \delta, \quad i = 1, \ldots, m
\]
\[
\|v\| \quad = \quad 1,
\]
\[v \in \mathbb{R}^n, \beta \in \mathbb{R}.\]

Now consider dividing each of the constraints of \(PCP\) by \(\delta\). If we perform the following transformation of variables:

\[x = \frac{v}{\delta}, \quad \alpha = \frac{\beta}{\delta}\]

then the constraints on the data points become:

\[x^T a^i - \alpha \quad \geq \quad 1, \quad i = 1, \ldots, k\]
\[
\alpha - x^T b^i \quad \geq \quad 1 \quad i = 1, \ldots, m.
\]

Also maximizing \(\delta\) is the same as maximizing \(\frac{\|v\|}{\|x\|} = \frac{1}{\|v\|}\), which is the same as minimizing \(\|x\|\). Therefore we can write the equivalent problem:

\[CPCP: \ \text{minimize}_{x, \alpha} \quad \|x\|\]
\[
\text{s.t.} \quad x^T a^i - \alpha \quad \geq \quad 1, \quad i = 1, \ldots, k
\]
\[
\alpha - x^T b^i \quad \geq \quad 1 \quad i = 1, \ldots, m
\]
\[x \in \mathbb{R}^n, \alpha \in \mathbb{R}.\]
The formulation \( CPCP \) has the nice property that all of the constraints are linear functions of the decision variables, and the objective function is the relatively simple function \( \|x\| \). We can solve \( CPCP \) for \( x, \alpha \), and substitute 
\[
v = \frac{x}{\|x\|}, \quad \beta = \frac{\alpha}{\|x\|} \quad \text{and} \quad \delta = \frac{1}{\|x\|}
\]
to obtain the solution of \( PCP \).

1.3.3 The Minimum Norm Problem

Given a vector \( c \), we would like to find the closest point to \( c \) that satisfies the linear inequalities \( Ax \leq b \).

This problem is formulated as:

\[
MNP : \quad \text{minimize}_{x} \quad \|x - c\|
\]

\[
\text{s.t.} \quad Ax \leq b
\]

\[
x \in \mathbb{R}^n.
\]

Figure 2 illustrates the formulation of the minimum norm problem.
1.3.4 The Fermat-Weber Problem

We are given \( m \) points \( c^1, \ldots, c^m \in \mathbb{R}^n \). We would like to determine the location of a distribution center at the point \( x \in \mathbb{R}^n \) that minimizes the sum of the distances from \( x \) to each of the points \( c^1, \ldots, c^m \in \mathbb{R}^n \). This problem is illustrated in Figure 3. It has the following formulation:

\[
FWP : \quad \text{minimize}_x \sum_{i=1}^{m} \|x - c^i\| \\
\text{s.t.} \quad x \in \mathbb{R}^n.
\]

1.3.5 A Ball Covering Problem

We are given \( m \) points \( c^1, \ldots, c^m \in \mathbb{R}^n \). We would like to determine the location of a distribution center at the point \( x \in \mathbb{R}^n \) that minimizes the maximum distance from \( x \) to any of the points \( c^1, \ldots, c^m \in \mathbb{R}^n \). This problem is illustrated in Figure 4. It has the following formulation:

\[
BCP : \quad \text{minimize}_{x, \delta} \quad \delta \\
\text{s.t.} \quad \|x - c^i\| \leq \delta, \quad i = 1, \ldots, m, \\
x \in \mathbb{R}^n.
\]

1.3.6 The Analytic Center Problem

Given a system of linear inequalities \( Ax \leq b \), we would like to determine a “nicely” interior point \( \hat{x} \) that satisfies \( A\hat{x} < b \). Of course, we would like the
Figure 3: Illustration of the Fermat-Weber problem.

Figure 4: Illustration of the ball covering problem.
point to be as interior as possible, in some mathematically meaningful way. Consider the following problem, which is referred to as the “analytic center problem”:

$$ACP : \text{maximize}_x \prod_{i=1}^{m} (b - Ax)_i$$

s.t. \quad Ax \leq b, \quad x \in \mathbb{R}^n.$$

Let us call the solution \( \hat{x} \) to this problem the “analytic center” of the linear inequality system “\( Ax \leq b \)”.

The analytic center problem is illustrated in Figure 5. The analytic center \( \hat{x} \) has the following rather remarkable (and non-obvious) property which we state but do not prove:

**Remark 1.1** Suppose that \( \hat{x} \) solves the analytic center problem \( ACP \). Then for each \( i = 1, \ldots, m \), it holds that:

$$\quad (b - A\hat{x})_i \geq \frac{\max_{x : Ax \leq b} (b - Ax)_i}{m}.$$

If we assume that the strict linear inequality system “\( Ax < b \)” has a solution, then the \( ACP \) is equivalent to:
$$CACP : \text{ minimize}_x \sum_{i=1}^{m} - \ln((b - Ax)_i)$$

$$\text{s.t. } Ax < b,$$

$$x \in \mathbb{R}^n.$$ 

1.3.7 Least Squares Problems, Model Reconstruction, and Curve-Fitting

We are given \(m\) functions \(g_1(x), \ldots, g_m(x)\), where \(g_i(x) : \mathbb{R}^n \mapsto \mathbb{R}^{r_i}, \ i = 1, \ldots, m\), and we are interested in finding a value of \(x\) that minimizes the sum of the squares of norms of these functions. This problem is:

$$LS : \text{ minimize}_x \frac{1}{2} \|g(x)\|^2 := \frac{1}{2} \sum_{i=1}^{m} \|g_i(x)\|^2$$

$$\text{s.t. } x \in \mathbb{R}^n.$$ 

Examples include linear and nonlinear regression, model reconstruction (inverse problems), and curve-fitting (splines).

In model reconstruction, we are given input/output data \((y_i, z_i), \ i = 1, \ldots, m\), from a physical system, and we postulate the input/output relationship

$$z = h(x, y)$$

where \(x\) is a vector of unknown parameters, and the general form of \(h(\cdot, \cdot)\) is known and given. We thus seek the value of the parameter vector \(x\) that best matches the data. One way to model this is to solve:
\[ MR : \text{ minimize}_x \quad \frac{1}{2} \sum_{i=1}^{m} ||z_i - h(x, y_i)||^2 \]
\[
\text{ s.t. } x \in \mathbb{R}^n .
\]

As a more specific example, let us consider the problem of fitting points to a quartic polynomial. Here

\[ h(x, y) = x_0 + x_1 y + x_2 y^2 + x_3 y^3 + x_4 y^4 , \]

and \( x = (x_0, x_1, x_2, x_3, x_4) \) is the vector of unknown coefficients of the polynomial. Suppose we are given data \((y_i, z_i), \ i = 1, \ldots, m\). Then we seek to solve:

\[ CF : \text{ minimize}_x \quad \frac{1}{2} \sum_{i=1}^{m} |z_i - (x_0 + x_1 y_i + x_2 y_i^2 + x_3 y_i^3 + x_4 y_i^4)|^2 \]
\[
\text{ s.t. } x \in \mathbb{R}^n .
\]

1.3.8 The Circumscribed Ellipsoid Problem

We are given \( m \) points \( c^1, \ldots, c^m \in \mathbb{R}^n \). We would like to determine an ellipsoid of minimum volume that contains each of the points \( c^1, \ldots, c^m \in \mathbb{R}^n \). This problem is illustrated in Figure 6.

Before we show the formulation of this problem, first recall that a symmetric positive definite (SPD) matrix \( R \) and a given point \( z \) can be used to define an ellipsoid in \( \mathbb{R}^n \):

\[ E_{R,z} := \{ y \mid (y - z)^T R (y - z) \leq 1 \} . \]
Here $R$ is the shape matrix of the ellipsoid and $z$ is the center of the ellipsoid. Figure 7 shows an illustration of an ellipsoid.

One can prove that the volume of $E_{R,z}$ is proportional to $\sqrt{\det(R^{-1})}$.

Our problem is:
\[ MCP_1 : \text{minimize} \quad [\text{det}(R^{-1})]^{\frac{1}{2}} \]
\[ \text{R, z} \quad \text{s.t.} \quad c^i \in E_{R, z}, \quad i = 1, \ldots, k, \]
\[ R \text{ is SPD} . \]

Of course, minimizing \( \text{det}(R^{-1})^{\frac{1}{2}} \) is the same as minimizing \( \ln(\text{det}(R^{-1})^{\frac{1}{2}}) \), since the logarithm function is strictly increasing in its argument. Also,

\[ \ln(\text{det}(R^{-1})^{\frac{1}{2}}) = -\frac{1}{2} \ln(\text{det}(R)) , \]

and so our problem is equivalent to:

\[ MCP_2 : \text{minimize} \quad -\ln(\text{det}(R)) \]
\[ \text{R, z} \quad \text{s.t.} \quad (c^i - z)^T R (c^i - z) \leq 1, \quad i = 1, \ldots, k \]
\[ R \text{ is SPD} . \]

We now factor \( R = M^2 \) where \( M \) is SPD (that is, \( M \) is a square root of \( R \)), and then \( MCP_2 \) becomes:

\[ MCP_3 : \text{minimize} \quad -\ln(\text{det}(M^2)) \]
\[ M, z \quad \text{s.t.} \quad (c^i - z)^T M^T M (c^i - z) \leq 1, \quad i = 1, \ldots, k, \]
\[ M \text{ is SPD} . \]

which is the same as:
MCP_4: minimize \(-2 \ln(\det(M))\) \quad M, z 
\text{s.t.} \quad \|M(c_i - z)\| \leq 1, \quad i = 1, \ldots, k, 
\quad M \text{ is SPD}.

Next substitute \(y = Mz\) to obtain:

MCP_5: minimize \(-2 \ln(\det(M))\) \quad M, y 
\text{s.t.} \quad \|Mc_i - y\| \leq 1, \quad i = 1, \ldots, k, 
\quad M \text{ is SPD}.

We can recover \(R\) and \(z\) after solving MCP_5 by substituting \(R = M^2\) and \(z = M^{-1}y\).

1.3.9 Portfolio Management Optimization

Portfolio optimization models are used throughout the financial investment management sector. These are nonlinear models that are used to determine the composition of investment portfolios.

Investors prefer higher annual rates of return on investing to lower annual rates of return. Furthermore, investors prefer lower risk to higher risk. Portfolio optimization seeks to optimally trade off risk and return in investing.

We consider \(n\) assets, whose annual rates of return \(R_i\) are random variables, \(i = 1, \ldots, n\). The expected annual return of asset \(i\) is \(\mu_i\), \(i = 1, \ldots, n\), and so if we invest a fraction \(x_i\) of our investment dollar in asset \(i\), the expected return of the portfolio is:

\[
\sum_{i=1}^{n} \mu_i x_i = \mu^T x
\]
where of course the fractions $x_i$ must satisfy:

$$\sum_{i=1}^{n} x_i = e^T x = 1.0$$

and

$$x \geq 0.$$ 

Here we use the notation that $e$ is the vector of ones, namely $e = (1, 1, \ldots, 1)^T$.

The covariance of the rates of return of assets $i$ and $j$ is given as

$$Q_{ij} = \text{COV}(R_i, R_j).$$

We can think of the $Q_{ij}$ values as forming a matrix $Q$, whereby the variance of portfolio is then:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \text{COV}(R_i, R_j)x_i x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij} x_i x_j = x^T Q x.$$

The “risk” in the portfolio is the standard deviation of the portfolio:

$$\text{STDEV} = \sqrt{x^T Q x},$$

and the “reward” of the portfolio is the expected annual rate of return of the portfolio:

$$\text{RETURN} = \mu^T x.$$ 

Suppose that we would like to determine the fractional investment values $x_1, \ldots, x_n$ in order to maximize the return of the portfolio, subject to meeting some pre-specified target risk level. For example, we might want to ensure that the standard deviation of the portfolio is at most $13.0\%$. We can formulate the following nonlinear optimization model:

\[
\begin{align*}
\text{MAXIMIZE:} & \quad \text{RETURN} = \mu^T x \\
\text{s.t.} & \quad e^T x = 1
\end{align*}
\]
An alternative version of the basic portfolio optimization model is to determine the fractional investment values $x_1, \ldots, x_n$ in order to minimize the risk of the portfolio, subject to meeting a pre-specified target expected return level. For example, we might want to ensure that the expected return of the portfolio is at least 16.0%. We can formulate the following nonlinear optimization model:

\[
\begin{align*}
\text{MINIMIZE:} & \quad \text{RISK} = \sqrt{x^TQx} \\
\text{s.t.} & \quad \text{FSUM:} \quad e^T x = 1 \\
& \quad \text{RETURN:} \quad \mu^T x \geq 16.0 \\
& \quad \text{NONNEGATIVITY:} \quad x \geq 0.
\end{align*}
\]

Variations of this family of optimization models are used pervasively in asset management companies worldwide. Also, these models can be extended to yield the CAPM (Capital Asset Pricing Model). Finally, we point out the Nobel Prize in economics was awarded in 1990 to Merton Miller, William Sharpe, and Harry Markowitz for their work on portfolio theory and portfolio models (and the implications for asset pricing).

The family of models just considered are conceptually simple but are obviously a simplification of the real situation. In reality, modern portfolio optimization models incorporate many features including limited short-selling, transaction costs, effects of buys and sells on market prices themselves, turnover, liquidity, and volatility constraints, regulatory restrictions, multiple-periods, and robustness considerations that explicitly take account of the fact that the expected return data $\mu$ and the covariance data $Q$ are not known with uncertainty, and hence the models must use imperfect estimates of the true data values.
1.4 Two Important Formats for the General Nonlinear Optimization Problem

There are two important formats for general nonlinear optimization problems. The first is the constraint functions format, which is as follows:

\[
\text{NLP} : \min_x f(x) \\
\text{s.t.} \quad g_1(x) \leq 0 \\
\vdots \\
g_m(x) \leq 0 \\
h_1(x) = 0 \\
\vdots \\
h_l(x) = 0 \\
x \in X ;
\]

where \(X\) is most typically the entire space \(X = \mathbb{R}^n\), but occasionally is some other open (or occasionally closed) set. Also,

\[
f(\cdot) : \mathbb{R}^n \mapsto \mathbb{R} ,
g_i(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}, i = 1, \ldots, m \\
h_j(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}, j = 1, \ldots, l
\]

where all of these functions are continuous and usually differentiable.

The feasible region of this problem is the set:

\[
\mathcal{F} = \{x : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, h_1(x) = 0, \ldots, h_l(x) = 0, \ x \in X\} .
\]
Let \( g(x) = (g_1(x), \ldots, g_m(x)) : \mathbb{R}^n \to \mathbb{R}^m \), \( h(x) = (h_1(x), \ldots, h_l(x)) : \mathbb{R}^n \to \mathbb{R}^l \). Then (NLP) can be written as
\[
\text{(NLP)} \quad \min_x \quad f(x) \\
\text{s.t.} \quad g(x) \leq 0 \\
\quad \quad \quad h(x) = 0 \\
\quad \quad \quad x \in X.
\] 

The other important format is the *subset inclusion* format, which is as follows:
\[
\text{(CNLS)} \quad \min_x \quad f(x) \\
\text{s.t.} \quad b - Ax \in K \\
\quad \quad \quad x \in C,
\] 
where the affine transformation \( x \mapsto b - Ax \) maps from \( \mathbb{R}^n \) to \( \mathbb{R}^m \), and \( K \subset \mathbb{R}^m \) and \( C \subset \mathbb{R}^n \). In most of this course, when we use this format, \( K \) and \( C \) will be *convex cones*, which we will discuss in a few lectures from now.

Here we say that the *feasible region* of this problem is the set:
\[\mathcal{F} = \{ x : b - Ax \in K, \ x \in C \}.\]

### 1.5 Some Special Classes of Nonlinear Optimization Problems

There are some useful special classes of nonlinear optimization problems. Let \( X = \mathbb{R}^n \).

An *unconstrained* nonlinear optimization problem is a problem with for-
A **linearly constrained** nonlinear optimization problem is a problem with format:

\[(\text{LC}) \quad \min_x f(x) \]
\[
\text{s.t.} \quad Ax \leq b \\
\quad Mx = d \\
\quad x \in X.
\]

A **linear optimization problem** is a problem with format:

\[(\text{LP}) \quad \min_x c^T x \\
\text{s.t.} \quad Ax \leq b \\
\quad Mx = d \\
\quad x \in X.
\]

A **quadratic optimization problem** is a problem with format:

\[(\text{QP}) \quad \min_x c^T x + \frac{1}{2} x^T Q x \\
\text{s.t.} \quad Ax \leq b \\
\quad Mx = d \\
\quad x \in X.
\]
A quadratically constrained optimization problem is a problem with format:

\[(QC) \quad \min_x \ f(x) \]
\[\text{s.t.} \quad g_i(x) := \frac{1}{2} x^T Q_i x + (c_i)^T x - d_i \leq 0, \ i = 1, \ldots, m \]
\[Mx = d \]
\[x \in X. \]

A separable optimization problem is a problem with format:

\[(SP) \quad \min_x \ \sum_{j=1}^n f_j(x) \]
\[\text{s.t.} \quad g_i(x) := \sum_{j=1}^n g_{ij}(x_j) \leq 0, \ i = 1, \ldots, m \]
\[h_i(x) := \sum_{j=1}^n h_{ij}(x_j) = 0, \ i = 1, \ldots, l \]
\[x \in X. \]

A conic optimization problem is a problem with format:

\[(CP) \quad \min_x \ c^T x \]
\[\text{s.t.} \quad b - Ax \in K \]
\[x \in C, \]

where $K$ and $C$ are convex cones (to be defined later).

1.6 Characterizations of Optimal Solutions: a Preview

From both a theoretical and computational point of view, it is important to understand ways to characterize (or just partially characterize) an optimal solution of a nonlinear optimization problem.
Consider the following 2-dimensional nonlinear optimization problem in variables \((x, y)\):

\[
(P_{a,b}) : \min_{x,y} \sqrt{(x-a)^2 + (y-b)^2}
\]

s.t. \[
\sqrt{(x-8)^2 + (y-9)^2} \leq 7
\]

\[
x \leq 13
\]

\[
x \geq 2
\]

\[
x + y \leq 24,
\]

where we are given the point \((a, b)\) as part of the problem description. The feasible region \(F\) of \(P_{a,b}\) is the intersection of four sets: (i) the disc of radius 7 centered at the point \((8, 9)\), (ii) the halfspace of points whose \(x\)-coordinate is at most 13, (iii) the halfspace of points whose \(x\)-coordinate is at least 2, and (iv) the halfspace characterized by \(x + y \leq 24\). Notice that the objective function of \(P_{a,b}\) is simply the Euclidean distance between \((x, y)\) and the point \((a, b)\).

Suppose that \((a, b) = (16, 14)\). Then the solution of \(P_{(a,b)}\) is the point \((x, y)^* = (13, 11)\), which is a “corner” of the feasible region \(F\) and is where the two constraints \(x \leq 13\) and \(x + y \leq 24\) are satisfied at equality.

Now suppose that \((a, b) = (14, 14)\). Then the solution of \(P_{(a,b)}\) is the point \((x, y)^* = (12, 12)\), which lies on the boundary of the feasible region in \(F\) but is not at a corner. Indeed, the point \((12, 12)\) lies in the “middle” of the boundary line segment corresponding to where the constraint \(x + y \leq 24\) is satisfied at equality.

Next suppose that \((a, b) = (8, 8)\). Then the solution of \(P_{(a,b)}\) is the point \((x, y)^* = (8, 8)\), which lies in the interior of the feasible region \(F\).

These examples illustrate that unlike in the case of linear optimization, the optimal solution does not necessarily occur at a “corner” of the feasible region or even on the boundary of the feasible region. The optimal solution will be where the feasible region just touches the best iso-objective function line. This concept will be made more precise later on.
2 Unconstrained Optimization: Characterization of Optimal Solutions

2.1 Local, Global, and Strict Optima

Consider the following optimization problem over the set $\mathcal{F}$:

$$
P : \begin{cases} 
\min_x \\
\max_x 
\end{cases} f(x) 
\text{ s.t. } x \in \mathcal{F}.
$$

Here $\mathcal{F}$ is the “feasible region” of the problem, and recall that $\mathcal{F}$ corresponds to $\mathcal{F} = \{x : g_1(x) \leq 0, \ldots, g_m(x) \leq 0, h_1(x) = 0, \ldots, h_l(x) = 0, x \in X\}$ for the constraint functions format for NLP in (2), or to $\mathcal{F} = \{x : b - Ax \in K, x \in C\}$ for the subset inclusion format (3). Of course, if there are no constraints in the problem, then $\mathcal{F} = \mathbb{R}^n$ and we will call $P$ an unconstrained optimization problem.

The ball centered at $\bar{x}$ with radius $\varepsilon$ is the set:

$$B(\bar{x}, \varepsilon) := \{x \|x - \bar{x}\| \leq \varepsilon\}.$$

We have the following definitions of local/global, strict/non-strict minima/maxima.

**Definition 2.1** $x \in \mathcal{F}$ is a local minimum of $P$ if there exists $\varepsilon > 0$ such that $f(x) \leq f(y)$ for all $y \in B(x, \varepsilon) \cap \mathcal{F}$.

**Definition 2.2** $x \in \mathcal{F}$ is a strict local minimum of $P$ if there exists $\varepsilon > 0$ such that $f(x) < f(y)$ for all $y \in B(x, \varepsilon) \cap \mathcal{F}, y \neq x$.

**Definition 2.3** $x \in \mathcal{F}$ is a global minimum of $P$ if $f(x) \leq f(y)$ for all $y \in \mathcal{F}$.

**Definition 2.4** $x \in \mathcal{F}$ is a strict global minimum of $P$ if $f(x) < f(y)$ for all $y \in \mathcal{F}, y \neq x$. 
Definition 2.5 \( x \in \mathcal{F} \) is a local maximum of \( P \) if there exists \( \varepsilon > 0 \) such that \( f(x) \geq f(y) \) for all \( y \in B(x, \varepsilon) \cap \mathcal{F} \).

Definition 2.6 \( x \in \mathcal{F} \) is a strict local maximum of \( P \) if there exists \( \varepsilon > 0 \) such that \( f(x) > f(y) \) for all \( y \in B(x, \varepsilon) \cap \mathcal{F}, y \neq x \).

Definition 2.7 \( x \in \mathcal{F} \) is a global maximum of \( P \) if \( f(x) \geq f(y) \) for all \( y \in \mathcal{F} \).

Definition 2.8 \( x \in \mathcal{F} \) is a strict global maximum of \( P \) if \( f(x) > f(y) \) for all \( y \in \mathcal{F}, y \neq x \).

2.2 Gradients and Hessians

Let \( f(\cdot) : X \rightarrow \mathbb{R} \), where \( X \subset \mathbb{R}^n \) is open. \( f(\cdot) \) is differentiable at \( \bar{x} \in X \) if there exists a vector \( \nabla f(\bar{x}) \) (called the gradient of \( f(\cdot) \) at \( \bar{x} \)) such that for each \( x \in X \) it holds that:

\[
f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \| x - \bar{x} \| \alpha(\bar{x}, x - \bar{x}),
\]

and \( \lim_{y \to 0} \alpha(\bar{x}, y) = 0 \). \( f(\cdot) \) is differentiable on \( X \) if \( f(x) \) is differentiable for all \( \bar{x} \in X \). In the standard coordinate system for \( \mathbb{R}^n \), the gradient vector corresponds to the vector of partial derivatives:

\[
\nabla f(\bar{x}) = \left( \frac{\partial f(\bar{x})}{\partial x_1}, \ldots, \frac{\partial f(\bar{x})}{\partial x_n} \right)^T.
\]

Example 2.1 Let \( f(x) = 3(x_1)^2(x_2)^3 + (x_2)^2(x_3)^3 \). Then

\[
\nabla f(x) = (6(x_1)(x_2)^2, 9(x_1)^2(x_2)^2 + 2(x_2)(x_3)^3, 3(x_2)^2(x_3)^2)^T.
\]

The directional derivative of \( f(\cdot) \) at \( \bar{x} \) in the direction \( d \) is:

\[
\lim_{\lambda \to 0} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}.
\]
Proposition 2.1 The directional derivative satisfies:

$$\lim_{\lambda \to 0} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^T d.$$ 

**Proof:** From the definition of the gradient, we have:

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \nabla f(\bar{x})^T (\lambda d) + \|\lambda d\| \alpha(\bar{x}, \lambda d),$$

where $\alpha(\bar{x}, y) \to 0$ as $y \to 0$. Rearranging and dividing by $\lambda$ yields:

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^T d + \|d\| \alpha(\bar{x}, \lambda d).$$

Now observe that the rightmost term goes to zero as $\lambda \to 0$, yielding the result. ■

The function $f(\cdot)$ is **twice differentiable** at $\bar{x} \in X$ if there exists a vector $\nabla f(\bar{x})$ and an $n \times n$ symmetric matrix $H(\bar{x})$ (called the **Hessian** of $f(\cdot)$ at $\bar{x}$) such that for each $x \in X$ it holds that:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T H(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}, x - \bar{x}),$$

and $\lim_{y \to 0} \alpha(\bar{x}, y) = 0$. $f(\cdot)$ is **twice differentiable on** $X$ if $f(x)$ is twice differentiable for all $x \in X$. In the standard coordinate system for $\mathbb{R}^n$, the Hessian corresponds to the matrix of second partial derivatives:

$$H(\bar{x})_{ij} = \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}.$$ 

Example 2.2 Continuing Example 2.1, we have

$$H(x) = \begin{pmatrix} 6(x_2)^3 & 18(x_1)(x_2)^2 & 0 \\ 18(x_1)(x_2)^2 & 18(x_1)^2(x_2) + 2(x_3)^3 & 6(x_2)(x_3)^2 \\ 0 & 6(x_2)(x_3)^2 & 6(x_3)^2 \end{pmatrix}. $$
2.3 Positive Semidefinite and Positive Definite Matrices

An $n \times n$ matrix $M$ is called:

- **positive definite** if $x^T M x > 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$
- **positive semidefinite** if $x^T M x \geq 0$ for all $x \in \mathbb{R}^n$
- **negative definite** if $x^T M x < 0$ for all $x \in \mathbb{R}^n$, $x \neq 0$
- **negative semidefinite** if $x^T M x \leq 0$ for all $x \in \mathbb{R}^n$
- **indefinite** if there exists $x, y \in \mathbb{R}^n$ for which $x^T M x > 0$ and $y^T M y < 0$.

We say that $M$ is SPD if $M$ is symmetric and positive definite. Similarly, we say that $M$ is SPSD if $M$ is symmetric and positive semidefinite.

**Example 2.3**

$$M = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

is positive definite.

**Example 2.4**

$$M = \begin{pmatrix} 8 & -1 \\ -1 & 1 \end{pmatrix}$$

is positive definite. To see this, note that for $x \neq 0$,

$$x^T M x = 8x_1^2 - 2x_1x_2 + x_2^2 = 7x_1^2 + (x_1 - x_2)^2 > 0 .$$

2.4 The Weierstrass Theorem, and Existence of Optimal Solutions

In order to study optimal solutions of optimization problems, we first need to know when an optimal solution exists. We will make use of the following important result from real analysis.
Theorem 2.1 [Weierstrass’ Theorem for sequences] Let \( \{x_k\} \), \( k \to \infty \) be an infinite sequence of points in the compact (i.e., closed and bounded) set \( \mathcal{F} \). Then some infinite subsequence of points \( x_{k_j} \) converges to a point contained in \( \mathcal{F} \).

Theorem 2.1 allows us to prove:

Theorem 2.2 [Weierstrass’ Theorem for functions] Let \( f(\cdot) \) be a continuous real-valued function on the compact nonempty set \( \mathcal{F} \subset \mathbb{R}^n \). Then \( \mathcal{F} \) contains a point that minimizes (maximizes) \( f(x) \) on the set \( \mathcal{F} \).

Proof: Since the set \( \mathcal{F} \) is non-empty and bounded, \( f(x) \) is bounded below on \( \mathcal{F} \). Therefore there exists \( v = \inf_{x \in \mathcal{F}} f(x) \). By definition, for any \( \varepsilon > 0 \), the set \( \mathcal{F}_\varepsilon = \{x \in \mathcal{F} : v \leq f(x) \leq v + \varepsilon\} \) is non-empty. Let \( \varepsilon_k \to 0 \) as \( k \to \infty \), and let \( x_k \in \mathcal{F}_{\varepsilon_k} \). Since \( \mathcal{F} \) is bounded, there exists a subsequence of \( \{x_k\} \) converging to some \( \bar{x} \in \mathcal{F} \). By continuity of \( f(\cdot) \), we have \( f(\bar{x}) = \lim_{k \to \infty} f(x_k) \), and, since \( v \leq f(x_k) \leq v + \varepsilon_k \), it follows that \( f(\bar{x}) = \lim_{k \to \infty} f(x_k) = v \).

Notice that the assumptions underlying Theorem 2.2 are (i) boundedness of the region \( \mathcal{F} \), (ii) closedness of the region \( \mathcal{F} \), and (iii) continuity of the function \( f(\cdot) \). Let us see why each of these conditions is needed. Consider the problem:

\[
(P) \quad \min_x \quad \frac{1 + x}{2x} \\
\text{s.t.} \quad x \geq 1 .
\]

Here there is no optimal solution because the feasible region \( \mathcal{F} \) is not bounded.

Next, consider the problem:

\[
(P) \quad \min_x \quad \frac{1}{x} \\
\text{s.t.} \quad 1 \leq x < 2 .
\]

Here there is no optimal solution because the feasible region \( \mathcal{F} \) is not closed.

Last of all, consider the problem:
\[
\begin{align*}
(P) \quad & \min_x \quad f(x) \\
\text{s.t.} \quad & 1 \leq x \leq 2 ,
\end{align*}
\]

where

\[
f(x) := \begin{cases} 
1/x & \text{for } 1 \leq x < 2 \\
1 & \text{for } x = 2 .
\end{cases}
\]

Here there is no optimal solution because the function \( f(\cdot) \) is not continuous.

It turns out that the assumptions of Theorem 2.2 can be somewhat relaxed.

For example, the following theorem can be proved in a manner similar to Theorem 2.2:

**Theorem 2.3** Consider the optimization problem:

\[
P: \min_x \quad f(x) \\
\text{s.t.} \quad x \in \mathcal{F} .
\]

Suppose that \( f(\cdot) \) is lower semi-continuous, i.e., for any constant \( c \), the set \( \{ x \in \mathcal{F} : f(x) \leq c \} \) is closed. Also, suppose that there exists some \( \hat{x} \in \mathcal{F} \) for which the set \( \{ x \in \mathcal{F} : f(x) \leq f(\hat{x}) \} \) is compact. Then \( \mathcal{F} \) contains a point that minimizes \( f(x) \) on the set \( \mathcal{F} \).

**2.5 Optimality Conditions for Unconstrained Problems**

Consider the unconstrained optimization problem:

\[
(P) \quad \min_x \quad f(x) \\
\text{s.t.} \quad x \in X ,
\]

where \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \), \( f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R} \), and \( X = \mathbb{R}^n \). Here the feasible region is the entire space \( X = \mathbb{R}^n \), so all points \( x \) are feasible.
Definition 2.9 The direction $\bar{d}$ is called a descent direction of $f(\cdot)$ at $x = \bar{x}$ if
$$f(\bar{x} + \varepsilon \bar{d}) < f(\bar{x})$$ for all $\varepsilon > 0$ and sufficiently small.

Proposition 2.2 Suppose that $f(\cdot)$ is differentiable at $\bar{x}$. If there is a vector $d$ such that $\nabla f(\bar{x})^T d < 0$, then $d$ is a descent direction of $f(\cdot)$ at $\bar{x}$. That is, $f(\bar{x} + \lambda d) < f(\bar{x})$ for all $\lambda > 0$ and sufficiently small.

Proof: From Proposition 2.1 we have:
$$\lim_{\lambda \to 0} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^T d.$$ Since $\nabla f(\bar{x})^T d < 0$, it follows that $f(\bar{x} + \lambda d) - f(\bar{x}) < 0$ for all $\lambda > 0$ sufficiently small. Therefore $d$ is a descent direction.

A necessary condition for local optimality is a statement of the form: “if $\bar{x}$ is a local minimum of $(P)$, then $\bar{x}$ must satisfy . . . ” Such a condition helps us identify all candidates for local optima. The following corollary provides a necessary condition for a point $\bar{x}$ to be a local minimum.

Corollary 2.1 Suppose $f(\cdot)$ is differentiable at $\bar{x}$. If $\bar{x}$ is a local minimum, then $\nabla f(\bar{x}) = 0$.

Proof: If it were true that $\nabla f(\bar{x}) \neq 0$, then $d = -\nabla f(\bar{x})$ would be a descent direction since then $\nabla f(\bar{x})^T d = -\nabla f(\bar{x})^T \nabla f(\bar{x}) = -\|\nabla f(\bar{x})\|^2 < 0$, whereby $\bar{x}$ would not be a local minimum.

The above corollary is a first order necessary optimality condition for an unconstrained minimization problem. The following theorem is a second order necessary optimality condition.

Theorem 2.4 Suppose that $f(\cdot)$ is twice continuously differentiable at $\bar{x} \in X$. If $\bar{x}$ is a local minimum, then $\nabla f(\bar{x}) = 0$ and $H(\bar{x})$ is positive semidefinite.
Proof: From the first order necessary condition in Corollary 2.1, it holds that $\nabla f(\bar{x}) = 0$. Suppose $H(\bar{x})$ is not positive semidefinite. Then there exists $d$ such that $d^T H(\bar{x})d < 0$. We have:

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^T d + \frac{1}{2} \lambda^2 d^T H(\bar{x})d + \lambda^2 \|d\|^2 \alpha(\bar{x}, \lambda d),$$

where $\alpha(\bar{x}, \lambda d) \to 0$ as $\lambda \to 0$. Rearranging,

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda^2} = \frac{1}{2} d^T H(\bar{x})d + \|d\|^2 \alpha(\bar{x}, \lambda d).$$

Since $d^T H(\bar{x})d < 0$ and $\alpha(\bar{x}, \lambda d) \to 0$ as $\lambda \to 0$, $f(\bar{x} + \lambda d) - f(\bar{x}) < 0$ for all $\lambda > 0$ sufficiently small, yielding the desired contradiction.

Example 2.5 Let

$$f(x) = \frac{1}{2} x_1^2 + x_1 x_2 + 2 x_2^2 - 4 x_1 - 4 x_2 - x_2^3.$$ 

Then

$$\nabla f(x) = \begin{pmatrix} x_1 + x_2 - 4, x_1 + 4 x_2 - 4 - 3 x_2^2 \end{pmatrix}^T,$$

and

$$H(x) = \begin{pmatrix} 1 & 1 \\ 1 & -6 x_2 \end{pmatrix}.$$ 

$\nabla f(x) = 0$ has exactly two solutions: $\bar{x} = (4, 0)$ and $\tilde{x} = (3, 1)$. We have

$$H(\bar{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$ 

and

$$H(\tilde{x}) = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}.$$ 

Notice that $H(\bar{x})$ is indefinite, whereby from Theorem 2.4 it follows that $\tilde{x} = (3, 1)$ cannot be a local minimum. Therefore the only possible candidate for a local minimum is $\bar{x} = (4, 0)$.

A sufficient condition for local optimality is a statement of the form: “if $\bar{x}$ satisfies . . . , then $\bar{x}$ is a local minimum of (P).” Such a condition allows us to automatically declare that $\bar{x}$ is indeed a local minimum.
Theorem 2.5 Suppose that \( f(\cdot) \) is twice differentiable at \( \bar{x} \). If \( \nabla f(\bar{x}) = 0 \) and \( H(\bar{x}) \) is positive definite, then \( \bar{x} \) is a (strict) local minimum.

Proof:

\[
f(x) = f(\bar{x}) + \frac{1}{2}(x - \bar{x})^T H(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}, x - \bar{x}).
\]

Suppose that \( \bar{x} \) is not a strict local minimum. Then there exists a sequence \( x_k \to \bar{x} \) such that \( x_k \neq \bar{x} \) and \( f(x_k) \leq f(\bar{x}) \) for all \( k \). Define \( d_k = \frac{\bar{x} - x_k}{\|\bar{x} - x_k\|} \), and notice that \( \|d_k\| = 1 \). Then

\[
f(x_k) = f(\bar{x}) + \|x_k - \bar{x}\|^2 \left( \frac{1}{2}d_k^T H(\bar{x})d_k + \alpha(\bar{x}, x_k - \bar{x}) \right),
\]

and so

\[
\frac{1}{2}d_k^T H(\bar{x})d_k + \alpha(\bar{x}, x_k - \bar{x}) = \frac{f(x_k) - f(\bar{x})}{\|x_k - \bar{x}\|^2} \leq 0.
\]

Since \( \|d_k\| = 1 \) for any \( k \), there exists a subsequence of \( \{d_k\} \) converging to some point \( d \) such that \( \|d\| = 1 \). Assume without loss of generality that \( d_k \to d \). Then

\[
0 \geq \lim_{k \to \infty} \left\{ \frac{1}{2}d_k^T H(\bar{x})d_k + \alpha(\bar{x}, x_k - \bar{x}) \right\} = \frac{1}{2}d^T H(\bar{x})d,
\]

which is a contradiction of the positive definiteness of \( H(\bar{x}) \).

Remark 2.1 If \( \nabla f(\bar{x}) = 0 \) and \( H(\bar{x}) \) is positive semidefinite, we cannot be sure if \( \bar{x} \) is a local minimum.

Remark 2.2 Similar to Theorem 2.5 it is easy to show that if \( \nabla f(\bar{x}) = 0 \) and \( H(\bar{x}) \) is negative definite, then \( \bar{x} \) is a (strict) local maximum.

Example 2.6 Continuing Example 2.5, we compute \( H(\bar{x}) \):

\[
H(\bar{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}.
\]

Then \( H(\bar{x}) \) is positive definite. To see this, note that for any \( d = (d_1, d_2) \), we have

\[
d^T H(\bar{x})d = d_1^2 + 2d_1d_2 + 4d_2^2 = (d_1 + d_2)^2 + 3d_2^2 > 0 \text{ for all } d \neq 0.
\]

Therefore, \( \bar{x} \) satisfies the sufficient conditions to be a strict local minimum, and so \( \bar{x} \) is a strict local minimum.
Example 2.7 Let \( f(x) = x_1^3 + x_2^2 \).

Then
\[
\nabla f(x) = (3x_1^2, 2x_2)^T,
\]

and
\[
H(x) = \begin{pmatrix} 6x_1 & 0 \\ 0 & 2 \end{pmatrix}.
\]

At \( \bar{x} = (0, 0) \), we have \( \nabla f(\bar{x}) = 0 \) and
\[
H(\bar{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}
\]
is positive semidefinite, but \( \bar{x} \) is not a local minimum, since \( f(-\varepsilon, 0) = -\varepsilon^3 < 0 = f(0, 0) = f(\bar{x}) \) for all \( \varepsilon > 0 \).

Example 2.8 Let \( f(x) = x_1^4 + x_2^2 \).

Then
\[
\nabla f(x) = (4x_1^3, 2x_2)^T,
\]

and
\[
H(x) = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{pmatrix}.
\]

At \( \bar{x} = (0, 0) \), we have \( \nabla f(\bar{x}) = 0 \) and
\[
H(\bar{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}
\]
is positive semidefinite. Furthermore, \( \bar{x} \) is a local minimum, since for all \( x \) we have \( f(x) \geq 0 = f(0, 0) = f(\bar{x}) \).

2.6 Convexity and Minimization

We now present some very important definitions regarding convexity of sets and functions.
Let \( x, y \in \mathbb{R}^n \). Points of the form \( \lambda x + (1 - \lambda)y \) for \( \lambda \in [0, 1] \) are called convex combinations of \( x \) and \( y \).

A set \( S \subset \mathbb{R}^n \) is called a convex set if for all \( x, y \in S \) and for all \( \lambda \in [0, 1] \) it holds that \( \lambda x + (1 - \lambda)y \in S \).

A function \( f(\cdot) : S \to \mathbb{R} \), where \( S \) is a nonempty convex set, is a convex function if

\[
f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)
\]

for all \( x, y \in S \) and for all \( \lambda \in [0, 1] \).

A function \( f(\cdot) \) as above is called a strictly convex function if the inequality above is strict for all \( x \neq y \) and \( \lambda \in (0, 1) \).

A function \( f(\cdot) : S \to \mathbb{R} \), where \( S \) is a nonempty convex set, is a concave function if

\[
f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)
\]

for all \( x, y \in S \) and for all \( \lambda \in [0, 1] \).

A function \( f(\cdot) \) as above is called a strictly concave function if the inequality above is strict for all \( x \neq y \) and \( \lambda \in (0, 1) \).

Consider the problem:

\[
(CP) \quad \min_x \quad f(x)
\]
\[\text{s.t.} \quad x \in S.
\]

**Theorem 2.6** Suppose \( S \) is a convex set, \( f(\cdot) : S \to \mathbb{R} \) is a convex function, and \( \bar{x} \) is a local minimum of (CP). Then \( \bar{x} \) is a global minimum of \( f(\cdot) \) over \( S \).

**Proof:** Suppose \( \bar{x} \) is not a global minimum. Then there exists \( y \in S \) for which \( f(y) < f(\bar{x}) \). Let \( y(\lambda) := (1 - \lambda)\bar{x} + \lambda y \) for \( \lambda \in [0, 1] \). Because \( y(\lambda) \) is a convex combination of \( \bar{x} \) and \( y \), it follows that \( y(\lambda) \in S \) because \( S \) is convex and both \( \bar{x} \) and \( y \) are elements of \( S \). Note that \( y(\lambda) \to \bar{x} \) as \( \lambda \to 0 \).
From the convexity of \( f(\cdot) \), it holds that:

\[
f(y(\lambda)) = f(\lambda y + (1-\lambda)\bar{x}) \leq \lambda f(y) + (1-\lambda)f(\bar{x}) < \lambda f(\bar{x}) + (1-\lambda)f(x) = f(\bar{x}),
\]

for all \( \lambda \in (0,1] \). Therefore, \( f(y(\lambda)) < f(\bar{x}) \) for all \( \lambda \in (0,1] \), and so \( \bar{x} \) is not a local minimum, resulting in a contradiction.

Similar to Theorem 2.6, one can prove:

**Corollary 2.2** The following are true:

- If \( f(\cdot) \) is strictly convex, a local minimum is the unique global minimum.
- If \( f(\cdot) \) is concave, a local maximum is a global maximum.
- If \( f(\cdot) \) is strictly concave, a local maximum is the unique global maximum.

The following result shows that a convex function always lies above its first-order approximation.

**Theorem 2.7** Suppose \( S \) is a non-empty open convex set, and \( f(\cdot) : S \to \mathbb{R} \) is differentiable. Then \( f(\cdot) \) is a convex function if and only if \( f(\cdot) \) satisfies the following gradient inequality:

\[
f(y) \geq f(x) + \nabla f(x)^T(y - x) \quad \text{for all } x, y \in S.
\]

**Proof:** Suppose \( f(\cdot) \) is convex. Then for any \( \lambda \in [0,1] \), it holds that:

\[
f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda)f(x).
\]

Now consider \( \lambda \in (0,1] \), and rearrange the above inequality to yield:

\[
\frac{f(x + \lambda(y-x)) - f(x)}{\lambda} \leq f(y) - f(x).
\]

Letting \( \lambda \to 0 \) and invoking Proposition 2.1, we obtain:

\[
\nabla f(x)^T(y - x) \leq f(y) - f(x),
\]

establishing the “only if” part.
Now suppose that the gradient inequality holds for all pairs of points in \( S \). Let \( w \) and \( z \) be any two points in \( S \). Let \( \lambda \in [0, 1] \), and set \( x = \lambda w + (1 - \lambda) z \). Then

\[
f(w) \geq f(x) + \nabla f(x)^T (w - x)
\]

and

\[
f(z) \geq f(x) + \nabla f(x)^T (z - x).
\]

Taking a convex combination of the above inequalities, we obtain

\[
\lambda f(w) + (1 - \lambda) f(z) \geq f(x) + \nabla f(x)^T (\lambda (w - x) + (1 - \lambda) (z - x))
\]

\[
= f(x) + \nabla f(x)^T 0
\]

\[
= f(\lambda w + (1 - \lambda) z),
\]

which shows that \( f(\cdot) \) is convex.

**Theorem 2.8** Suppose \( S \) is a non-empty open convex set, and \( f(\cdot) : S \to \mathbb{R} \) is twice differentiable. Let \( H(\cdot) \) denote the Hessian of \( f(\cdot) \). Then \( f(\cdot) \) is convex if and only if \( H(x) \) is positive semidefinite for all \( x \in S \).

**Proof:** Suppose \( f(\cdot) \) is convex. Let \( \bar{x} \in S \) and \( d \) be any direction. Then for \( \lambda > 0 \) sufficiently small, \( \bar{x} + \lambda d \in S \). We have:

\[
f(\bar{x} + \lambda d) = f(\bar{x}) + \nabla f(\bar{x})^T (\lambda d) + \frac{1}{2} (\lambda d)^T H(\bar{x})(\lambda d) + \|\lambda d\|^2 \alpha(\bar{x}, \lambda d),
\]

where \( \alpha(\bar{x}, y) \to 0 \) as \( y \to 0 \). From the gradient inequality, we also have

\[
f(\bar{x} + \lambda d) \geq f(\bar{x}) + \nabla f(\bar{x})^T (\lambda d).
\]

Combining the above two relations, we obtain:

\[
\lambda^2 \left( \frac{1}{2} d^T H(\bar{x}) d + \|d\|^2 \alpha(\bar{x}, \lambda d) \right) \geq 0.
\]
Dividing by $\lambda^2 > 0$ and letting $\lambda \to 0$, we obtain $d^T H(x)d \geq 0$, proving the "only if" part.

Conversely, suppose that $H(z)$ is positive semidefinite for all $z \in S$. Let $x, y \in S$ be given. Then according to the second-order version of the Mean Value Theorem, there exists a value $z$ which is a convex combination of $x$ and $y$ which satisfies the following equation:

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)^T H(z)(y - x).$$

Since $z$ is a convex combination of $x$ and $y$, then $z = \lambda x + (1 - \lambda)y$ for some $\lambda \in [0, 1]$, and hence $z \in S$. Therefore $H(z)$ is positive semidefinite, and this implies that

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

Therefore the gradient inequality holds, and hence $f(\cdot)$ is convex as a consequence of Theorem 2.7.

Returning to the optimization problem (P), knowing that the function $f(\cdot)$ is convex allows us to establish a global optimality condition that is both necessary and sufficient:

**Theorem 2.9** Suppose $f(\cdot) : X \to \mathbb{R}$ is convex and differentiable on $X$. Then $\bar{x} \in X$ is a global minimum if and only if $\nabla f(\bar{x}) = 0$.

**Proof:** The necessity of the condition $\nabla f(\bar{x}) = 0$ was established in Corollary 4 regardless of the convexity of the function $f(\cdot)$.

Suppose $\nabla f(\bar{x}) = 0$. Then by the gradient inequality we have $f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^T (y - \bar{x}) = f(\bar{x})$ for all $y \in X$, and so $\bar{x}$ is a global minimum.

**Example 2.9** Continuing Example 2.5, recall that

$$H(x) = \begin{pmatrix} 1 & 1 \\ 1 & 4 - 6x_2 \end{pmatrix}. $$

Suppose that the domain of $f(\cdot)$ is $X = \{(x_1, x_2) \mid x_2 < \frac{1}{2}\}$. Then $f(\cdot)$ is a convex function on this domain.
Example 2.10 Let
\[ f(x) = -\ln(1 - x_1 - x_2) - \ln x_1 - \ln x_2. \]

Then
\[ \nabla f(x) = \begin{bmatrix} \frac{1}{1-x_1-x_2} - \frac{1}{x_1} \\ \frac{1}{1-x_1-x_2} - \frac{1}{x_2} \end{bmatrix}, \]
and
\[ H(x) = \begin{bmatrix} \left(\frac{1}{1-x_1-x_2}\right)^2 + \left(\frac{1}{x_1}\right)^2 & \left(\frac{1}{1-x_1-x_2}\right)^2 \\ \left(\frac{1}{1-x_1-x_2}\right)^2 & \left(\frac{1}{1-x_1-x_2}\right)^2 + \left(\frac{1}{x_2}\right)^2 \end{bmatrix}. \]

It is actually easy to prove that \( f(\cdot) \) is a strictly convex function, and hence that \( H(x) \) is positive definite on its domain \( \mathcal{X} = \{(x_1, x_2) \mid x_1 > 0, x_2 > 0, x_1 + x_2 < 1\} \). At \( \bar{x} = \left(\frac{1}{3}, \frac{1}{3}\right) \) we have \( \nabla f(\bar{x}) = 0 \), and so \( \bar{x} \) is the unique global minimum of \( f(\cdot) \).

3 A Line-search Algorithm for a Convex Function based on Bisection

Let us consider the unconstrained optimization problem:
\[ \text{(P)} \quad \min_x f(x) \]
\[ \text{ s.t. } x \in \mathbb{R}^n. \]

In a typical algorithm for nonlinear optimization, the current iterate is some value \( \bar{x} \in \mathcal{F} \), and the algorithm then computes a direction \( \bar{d} \) that satisfies
\[ \nabla f(\bar{x})^T \bar{d} < 0, \quad (4) \]
whereby we know from Proposition 2.2 that \( \bar{d} \) is a descent direction of \( f(\cdot) \) at \( x = \bar{x} \). A typical algorithm then determines the next iterate \( \bar{x}^{\text{new}} \) by computing a step-size \( \alpha \) and setting:
\[ \bar{x}^{\text{new}} \leftarrow \bar{x} + \alpha \bar{d}. \]
It is often desirable to choose the step-size value $\alpha$ so that the objective function value of the new iterate is improved as much as possible, and therefore we wish to determine the optimal value $\alpha^*$ of the following 1-dimensional problem:

$$\alpha^* \leftarrow \arg \min_{\alpha \geq 0} f(\bar{x} + \alpha \bar{d}) .$$

Problem (5) is called a line-search problem, as we seek to search for the best value of $\alpha$ on the (half-) line $L = \{ \bar{x} + \alpha \bar{d} : \alpha \geq 0 \}$.

Let us consider the case when $f(\cdot)$ is a continuously differentiable convex function. Let

$$h(\alpha) := f(\bar{x} + \alpha \bar{d}) .$$

Then $h(\cdot)$ is a convex function in the scalar variable $\alpha$, and our problem then is to solve for

$$\alpha^* = \arg \min_{\alpha \geq 0} h(\alpha) .$$

Due to the convexity of $h(\cdot)$, we equivalently seek a value $\alpha^*$ for which

$$h'(\alpha^*) = 0 .$$

We have:

**Proposition 3.1** The following properties of $h(\cdot)$ hold:

1. $h'(\alpha) = \nabla f(\bar{x} + \alpha \bar{d})^T \bar{d}$,
2. $h'(0) < 0$, and
3. $h'(\alpha)$ is a monotone increasing function of $\alpha$.

**Proof:** Item (1.) follows from the definition of the directional derivative and Proposition 2.1, and item (2.) follows from (4). Item (3.) is a basic property of a 1-dimensional convex function, namely that its slope is non-decreasing.

Because $h'(\alpha)$ is a monotonically increasing function, we can approximately compute $\alpha^*$, the point that satisfies $h'(\alpha^*) = 0$, by a suitable bisection method. Let us see how this can be done. Suppose that we know a value $\hat{\alpha}$ that $h'(\hat{\alpha}) > 0$. Since $h'(0) < 0$ and $h'(\hat{\alpha}) > 0$, by the Mean Value Theorem
there exists some value of $\alpha \in (0, \tilde{\alpha})$ for which $h'(\alpha) = 0$; in other words we have $\alpha^* \in (0, \tilde{\alpha})$. We will use the mid-value $\tilde{\alpha} = \frac{0+\tilde{\alpha}}{2}$ as a suitable test-point. Note the following:

- If $h'(\tilde{\alpha}) = 0$, we are done.
- If $h'(\tilde{\alpha}) > 0$, then $\alpha^* \in (0, \tilde{\alpha})$.
- If $h'(\tilde{\alpha}) < 0$, then $\alpha^* \in (\tilde{\alpha}, \tilde{\alpha})$.

Repeating this process, we are led to the following bisection algorithm presented in Algorithm 1 for solving the line-search problem for approximately computing $\alpha^* = \arg \min_{\alpha \geq 0} h(\alpha) = \arg \min_{\alpha \geq 0} f(\bar{x} + \alpha d)$ by approximately solving the equation $h'(\alpha) = 0$.

**Algorithm 1** Bisection Algorithm for Line-search Problem

Initialize at $\alpha_l := 0$ and $\alpha_u := \tilde{\alpha}$. Set $k \leftarrow 0$.

At iteration $k$:

1. Set $\tilde{\alpha} = \frac{\alpha_u + \alpha_l}{2}$ and compute $h'(&tilde{\alpha})$.
   - If $h'(\tilde{\alpha}) > 0$, re-set $\alpha_u := \tilde{\alpha}$. Set $k \leftarrow k + 1$.
   - If $h'(\tilde{\alpha}) < 0$, re-set $\alpha_l := \tilde{\alpha}$. Set $k \leftarrow k + 1$.
   - If $h'(\tilde{\alpha}) = 0$, stop.

The following properties of the bisection algorithm for solving the line-search problem are easy to establish:

**Proposition 3.2**

1. After every iteration of the bisection algorithm, the current interval $[\alpha_l, \alpha_u]$ contains a point $\alpha^*$ for which $h'(\alpha^*) = 0$.

2. At the $k^{th}$ iteration of the bisection algorithm, the length of the current interval $[\alpha_l, \alpha_u]$ is

$$L = \left(\frac{1}{2}\right)^k (\tilde{\alpha})$$
3. A value of $\alpha$ such that $|\alpha - \alpha^*| \leq \varepsilon$ can be found in at most
\[
\left\lceil \log_2 \left( \frac{\hat{\alpha}}{\varepsilon} \right) \right\rceil
\]
steps of the bisection algorithm.

3.1 Computing $\hat{\alpha}$ for which $h'(\hat{\alpha}) > 0$

Suppose that we do not have available a convenient value $\hat{\alpha}$ for which $h'(\hat{\alpha}) > 0$. One way to proceed is to pick an initial “guess” of $\hat{\alpha} > 0$ and compute $h'(\hat{\alpha})$. If $h'(\hat{\alpha}) > 0$, then proceed with the bisection algorithm; if $h'(\hat{\alpha}) \leq 0$, then re-set $\hat{\alpha} \leftarrow 2\hat{\alpha}$ and repeat the process.

3.2 Stopping Criteria for the Bisection Algorithm

In practice, we need to run the bisection algorithm with a stopping criterion. Some relevant stopping criteria are:

- Stop after a fixed number of iterations. That is, stop when $k = \bar{K}$, where $\bar{K}$ is specified by the user.
- Stop when the interval becomes small. That is, stop when $\alpha_u - \alpha_l \leq \varepsilon$, where $\varepsilon$ is specified by the user.
- Stop when $|h'(\hat{\alpha})|$ becomes small. That is, stop when $|h'(\hat{\alpha})| \leq \varepsilon$, where $\varepsilon$ is specified by the user.

The third stopping criterion typically yields the best results in practice.
3.3 Modification of the Bisection Algorithm when the Domain of $f(\cdot)$ is Restricted

The discussion and analysis of the bisection algorithm has presumed that our optimization problem is

$$P: \text{minimize}_x \ f(x)$$

$$\text{s.t.} \quad x \in \mathbb{R}^n.$$ 

Given a point $\bar{x}$ and a descent direction $\bar{d}$, the line-search problem then is

$$LS: \text{minimize}_\alpha \ h(\alpha) := f(\bar{x} + \alpha \bar{d})$$

$$\text{s.t.} \quad \alpha \geq 0.$$ 

Suppose instead that the domain of definition of $f(\cdot)$ is an open set $\mathcal{F} \subset \mathbb{R}^n$. Then our optimization problem is:

$$P: \text{minimize}_x \ f(x)$$

$$\text{s.t.} \quad x \in \mathcal{F},$$

and the line-search problem then is

$$LS: \text{minimize}_\alpha \ h(\alpha) := f(\bar{x} + \alpha \bar{d})$$

$$\text{s.t.} \quad \bar{x} + \alpha \bar{d} \in \mathcal{F}$$

$$\alpha \geq 0.$$ 

In this case, we must ensure that all iterate values of $\alpha$ in the bisection
algorithm satisfy $\bar{x} + \alpha \bar{d} \in \mathcal{F}$. As an example, consider the following problem:

$$P: \quad \text{minimize}_{\bar{x}} \quad f(x) := -\sum_{i=1}^{m} \ln(b_i - A_i \bar{x})$$

s.t. \quad b - Ax > 0.

Here the domain of $f(\cdot)$ is $\mathcal{F} = \{x \in \mathbb{R}^n \mid b - Ax > 0\}$. Given a point $\bar{x} \in X$ and a descent direction $\bar{d}$, the line-search problem is:

$$LS: \quad \text{minimize}_{\alpha} \quad h(\alpha) := f(\bar{x} + \alpha \bar{d}) = -\sum_{i=1}^{m} \ln(b_i - A_i (\bar{x} + \alpha \bar{d}))$$

s.t. \quad b - A(\bar{x} + \alpha \bar{d}) > 0

$$\alpha \geq 0.$$ 

Standard arithmetic manipulation can be used to establish that

$$b - A(\bar{x} + \alpha \bar{d}) > 0 \quad \text{if} \quad 0 \leq \alpha < \hat{\alpha}$$

where

$$\hat{\alpha} := \min_{A_i \bar{d} > 0} \left\{ \frac{b_i - A_i \bar{x}}{A_i \bar{d}} \right\},$$

and the line-search problem then is:

$$LS: \quad \text{minimize}_{\alpha} \quad h(\alpha) := -\sum_{i=1}^{m} \ln(b_i - A_i (\bar{x} + \alpha \bar{d}))$$

s.t. \quad 0 \leq \alpha < \hat{\alpha}.$$

4 Exercises on Unconstrained Optimization

1. Find points satisfying necessary conditions for extrema (that is, local minima or local maxima) of the function

$$f(x) = \frac{x_1 + x_2}{3 + x_1^2 + x_2^2 + x_1 x_2}.$$
Try to establish the nature of these points by checking sufficient conditions.

2. Find minima of the function
\[ f(x) = (x_2^2 - x_1)^2 \]
among all the points satisfying necessary conditions for an extremum.

3. Consider the problem of finding \( x \) that minimizes \( ||Ax - b||^2 \), where \( A \) is an \( m \times n \) matrix and \( b \) is an \( m \)-dimensional vector.
   a. Give a geometric interpretation of the problem.
   b. Write a necessary condition for optimality. Is this also a sufficient condition?
   c. Is the optimal solution unique? Why or why not?
   d. Can you construct a closed-form expression for the optimal solution? Specify any assumptions that you may need.
   e. Solve the problem for \( A \) and \( b \) given below:

\[
A = \begin{pmatrix}
1 & -1 & 0 \\
0 & 2 & 1 \\
0 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}.
\]

4. Let \( S \) be a nonempty set in \( \mathbb{R}^n \). Show that \( S \) is convex if and only if for each integer \( k \geq 2 \) the following holds true:

\[
x^1, \ldots, x^k \in S \Rightarrow \sum_{j=1}^{k} \lambda_j x^j \in S
\]

whenever \( \lambda_1, \ldots, \lambda_k \) satisfy \( \lambda_1, \ldots, \lambda_k \geq 0 \) and \( \sum_{j=1}^{k} \lambda_j = 1 \).

5. Prove Corollary 2.2.

6. Consider the function \( f(\cdot) = f_a(\cdot) \) parameterized by the scalar \( a \):

\[
f_a(x_1, x_2) = 4(x_1)^2 + (x_2)^2 + a(x_1)(x_2) + 2x_1 + x_2.
\]

For each value of \( a \), describe the set of solutions of \( \nabla f_a(x) = 0 \). Which of these solutions are global minima?
7. Bertsekas, Exercise 1.1.2, page 16, parts (a), (b), (c), and (d). (Note: $x^*$ is called a **stationary point** of $f(\cdot)$ if $\nabla f(x^*) = 0$.)

8. Let $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\bar{x}$, and let $d_1, \ldots, d_n$ be linearly independent vectors in $\mathbb{R}^n$. For each $j = 1, \ldots, n$, suppose that the minimum of $f(\bar{x} + \lambda d_j)$ over $\lambda \in \mathbb{R}$ occurs at $\lambda = 0$. Show that this implies that $\nabla f(\bar{x}) = 0$. Does this imply that $f(\cdot)$ has a local minimum at $\bar{x}$?