Lecture 2: Optimization

Prof. Dr. Svetlozar Rachev

Institute for Statistics and Mathematical Economics
University of Karlsruhe

Portfolio and Asset Liability Management

Summer Semester 2008

Copyright

The material is based on the text-book:
Svetlozar T. Rachev, Stoyan Stoyanov, and Frank J. Fabozzi
Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures
John Wiley, Finance, 2007

Prof. Svetlozar (Zari) T. Rachev
Chair of Econometrics, Statistics
and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss, Bau II, 20.12, R210
Postfach 6980, D-76128, Karlsruhe, Germany
Tel. +49-721-608-7535, +49-721-608-2042(s)
Fax: +49-721-608-3811
http://www.statistik.uni-karslruhe.de

Basic concepts

- The portfolio choice problem concerns the optimal trade-off between risk and reward. An optimal portfolio is the best portfolio among many alternative ones.
- The criterion which measures the “quality” of a portfolio relative to the others is known as the objective function in optimization theory.
- The set of portfolios among which we are choosing is called the set of feasible solutions or the set of feasible points.

Basic concepts

- There are two types of optimization problems depending on whether the set of feasible solutions is constrained or unconstrained.
- If the optimization problem is a constrained one, then the set of feasible solutions is defined by means of certain linear and/or non-linear equalities and inequalities. These functions are forming the constraint set.
- There are also types of optimization problems depending on the assumed properties of the objective function and the functions in the constraint set, such as linear problems, quadratic problems, and convex problems.
- The solution methods vary with respect to the particular optimization problem type as there are efficient algorithms prepared for particular problem types.

Unconstrained optimization

- *Unconstrained optimization problem* is when there are no constraints imposed on the set of feasible solutions. Thus, the goal is to maximize or to minimize the objective function with respect to the function arguments without any limits on their values.
- Consider the n-dimensional case; that is, the domain of the objective function f is the n-dimensional space and the function values are real numbers, $f : \mathbb{R}^n \to \mathbb{R}$.
  Maximization is denoted by

$$\max f(x_1, \ldots, x_n)$$

  and minimization by

$$\min f(x_1, \ldots, x_n).$$

Unconstrained optimization

- A more compact form is commonly used, for example

$$\min_{x \in \mathbb{R}^n} f(x) \tag{1}$$

denotes that we are searching for the minimal value of the function $f(x)$ by varying x in the entire n-dimensional space $\mathbb{R}^n$.

- A solution to problem (1) is a value of $x = x^0$ for which the minimum of f is attained,

$$f_0 = f(x^0) = \min_{x \in \mathbb{R}^n} f(x).$$

- Thus, the vector x_0 is such that the function takes a larger value than f_0 for any other vector x,

$$f(x^0) \leq f(x), \; x \in \mathbb{R}^n. \tag{2}$$

Remark: there may be more than one vector x^0 and, therefore, the argument for which f_0 is achieved may not be unique.

Unconstrained optimization

- If (2) holds, then the function is attaining its global minimum at x^0.
- If the inequality in (2) holds for x belonging only to a small neighborhood of x^0 and not to the entire space $\mathbb{R}^n$, then the objective function is said to have a local minimum at x^0:

$$f(x^0) \leq f(x)$$

for all x such that $||x - x^0||_2 < \epsilon$ where $||x - x^0||_2$ stands for the Euclidean distance between the vectors x and x^0,

$$||x - x^0||_2 = \sqrt{\sum_{i=1}^{n} (x_i - x_i^0)^2},$$

and ϵ is some positive number.

- A local minimum may not be global as there may be vectors outside the small neighborhood of x_0 for which the objective function attains a smaller value than $f(x_0)$.

Unconstrained optimization

Figures below show the graph of a function with two local maxima, one of which is the global maximum.

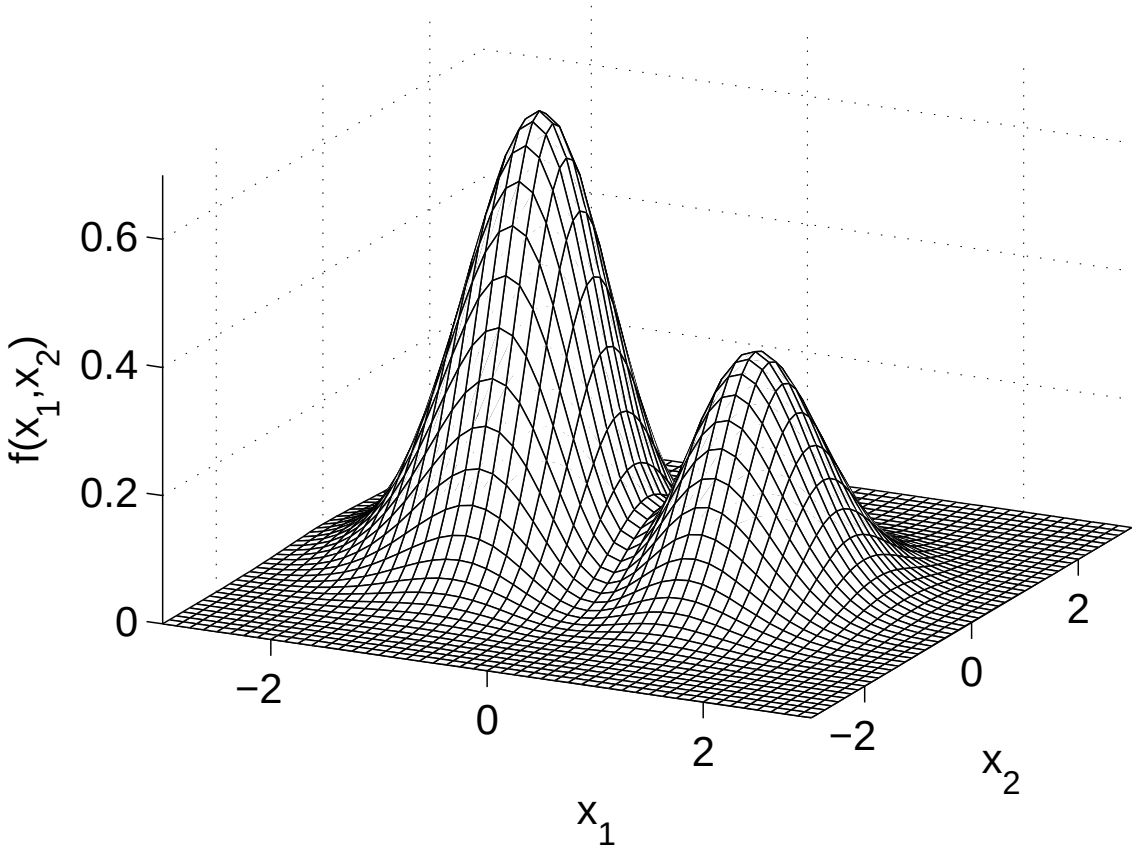

Figure: The plot shows a function $f(x_1, x_2)$ with two local maxima.

Unconstrained optimization

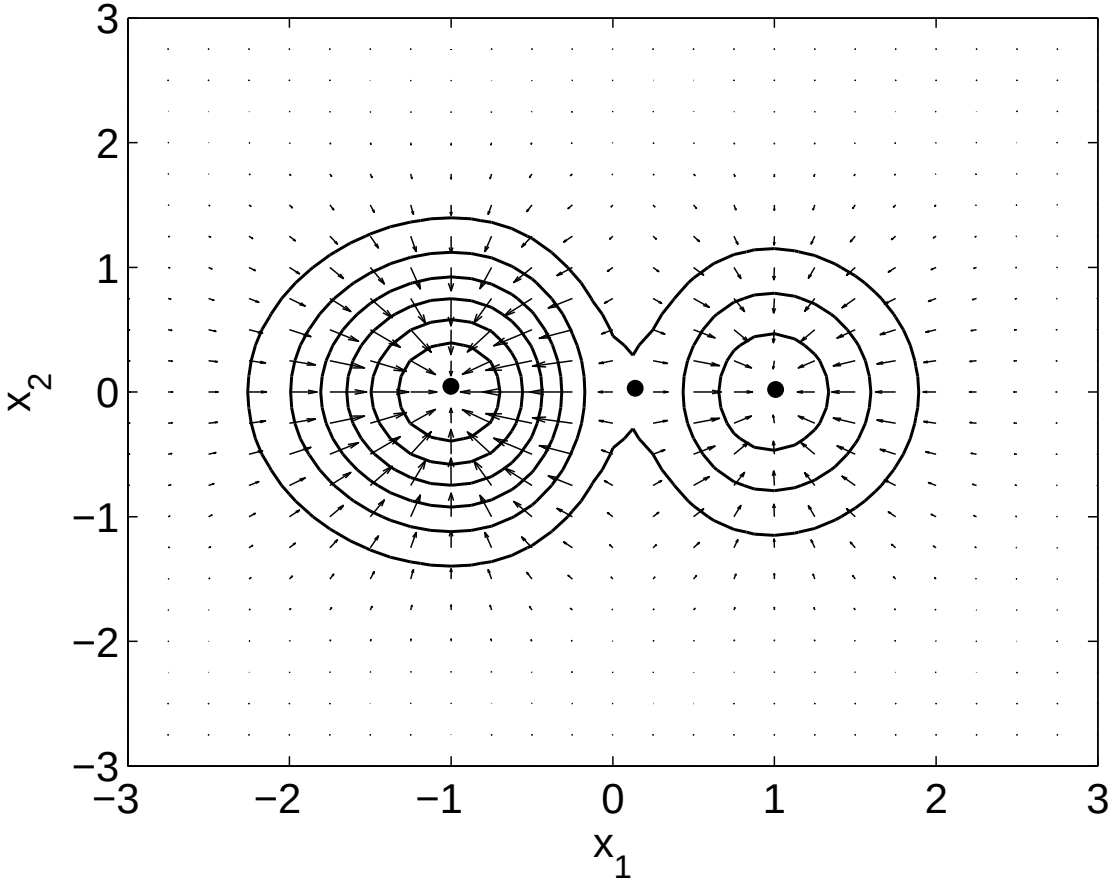

Figure: The plot shows the contour lines of $f(x_1, x_2)$ together with the gradient evaluated at a grid of points. The middle black point shows the position of the saddle point between the two local maxima.

Unconstrained optimization

- Maximizing the objective function is the same as minimizing the negative of the objective function and then changing the sign of the minimal value,

$$\max_{x \in \mathbb{R}^n} f(x) = -\min_{x \in \mathbb{R}^n} [-f(x)].$$

- This relationship between minimization and maximization is illustrated in the figure on the next slide. As a consequence, problems for maximization can be stated in terms of function minimization and vice versa.

Unconstrained optimization

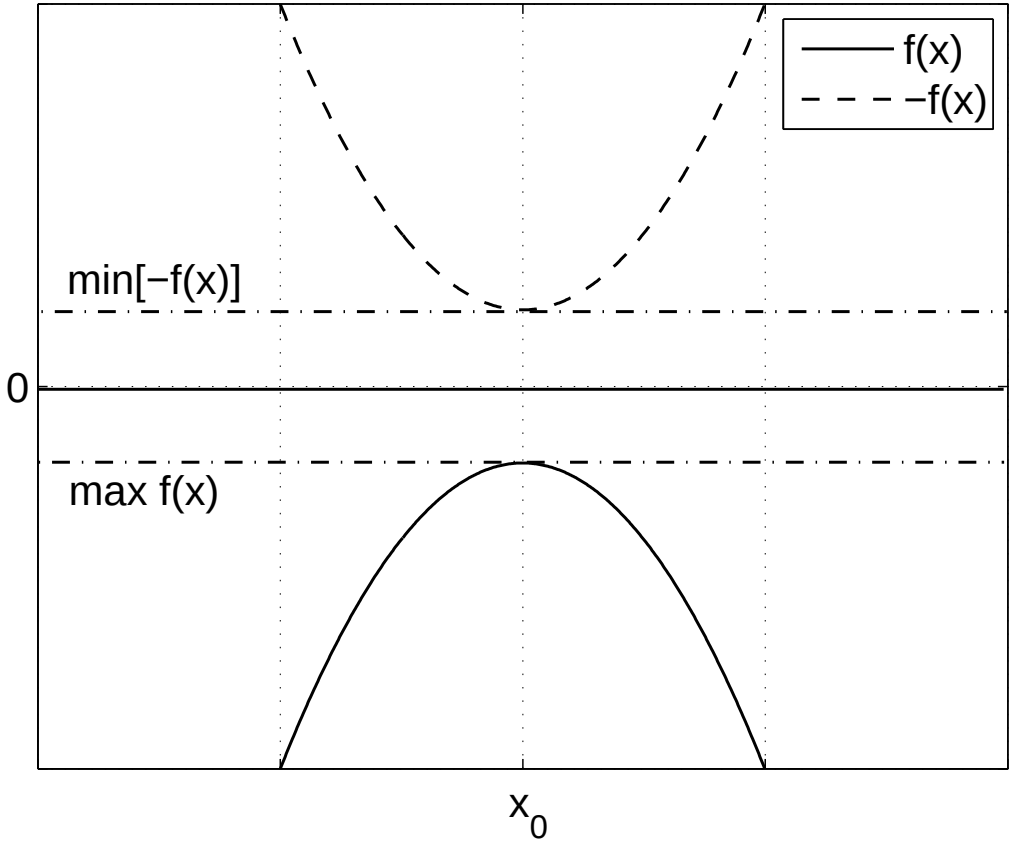

Figure: The relationship between minimization and maximization for a one-dimensional function.

Minima and maxima of a differentiable function

- If the second derivatives of the objective function exist, then its local maxima and minima, often called generically local extrema, can be characterized.
- Denote by $\nabla f(x)$ the vector of the first partial derivatives of the objective function evaluated at x,

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n} \right).$$

- This vector is called the function gradient. At each point x of the domain of the function, it shows the direction of greatest rate of increase of the function in a small neighborhood of x.

Minima and maxima of a differentiable function

- If for a given x, the gradient equals a vector of zeros,

$$\nabla f(x) = (0, \ldots, 0)$$

  then the function does not change in a small neighborhood of $x \in \mathbb{R}^n$.
- It turns out that all points of local extrema of the objective function are characterized by a zero gradient. As a result, the points yielding the local extrema of the objective function are among the solutions of the system of equations,

$$\left| \begin{array}{l} \dfrac{\partial f(x)}{\partial x_1} = 0 \\ \ldots \\ \dfrac{\partial f(x)}{\partial x_n} = 0. \end{array} \right. \tag{3}$$

Minima and maxima of a differentiable function

- The system of equations (3) is referred to as representing the first-order condition for the objective function extrema. However, it is only a necessary condition; that is, if the gradient is zero at a given point in the n-dimensional space, then this point may or may not be a point of a local extremum for the function.
- An illustration was given in the slides 8,9. The middle point, called saddle point, is not a point of a local maximum even though it has a zero gradient. It is called saddle since the graph resembles the shape of a saddle in a neighborhood of it.
- This example demonstrates that the first-order conditions are generally insufficient to characterize the points of local extrema.

Minima and maxima of a differentiable function

- The additional condition which identifies which of the zero-gradient points are points of local minimum or maximum is given through the matrix of second derivatives,

$$H = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}, \tag{4}$$

  which is called the Hessian matrix or just the Hessian.
- The Hessian is a symmetric matrix because the order of differentiation is insignificant,

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

Minima and maxima of a differentiable function

The additional condition is known as the second-order condition. Second-order condition for functions of n-dimensional arguments is rather technical, so we only state it for two-dimensional functions. In the case $n = 2$, the following conditions hold:

- If $\nabla f(x_1, x_2) = (0, 0)$ at a given point (x_1, x_2) and the determinant of the Hessian matrix evaluated at (x_1, x_2) is positive, then the function has

  a local maximum in (x_1, x_2) if

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} < 0 \quad \text{or} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} < 0$$

  a local minimum in (x_1, x_2) if

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} > 0 \quad \text{or} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} > 0$$

Minima and maxima of a differentiable function

- If $\nabla f(x_1, x_2) = (0, 0)$ at a given point (x_1, x_2) and the determinant of the Hessian matrix evaluated at (x_1, x_2) is negative, then the function f has a saddle point in (x_1, x_2)
- If $\nabla f(x_1, x_2) = (0, 0)$ at a given point (x_1, x_2) and the determinant of the Hessian matrix evaluated at (x_1, x_2) is zero, then no conclusion can be drawn.

Convex functions

- We showed that the first-order conditions are insufficient in the general case to describe the local extrema. However, when certain assumptions are made for the objective function, the first-order conditions can become sufficient.
- Furthermore, for certain classes of functions, the local extrema are necessarily global. Therefore, solving the first-order conditions we obtain the global extremum.
- A general class of functions with nice optimal properties is the class of convex functions.

Convex functions

- Precisely, a function $f(x)$ is called a convex function if it satisfies the property: For a given $\alpha \in [0, 1]$ and all $x^1 \in \mathbb{R}^n$ and $x^2 \in \mathbb{R}^n$ in the function domain,

$$f(\alpha x^1 + (1 - \alpha)x^2) \leq \alpha f(x^1) + (1 - \alpha)f(x^2). \tag{5}$$

- The definition is illustrated in Figure on the next slide. Basically, if a function is convex, then a straight line connecting any two points on the graph lies "above" the graph of the function.

Convex functions

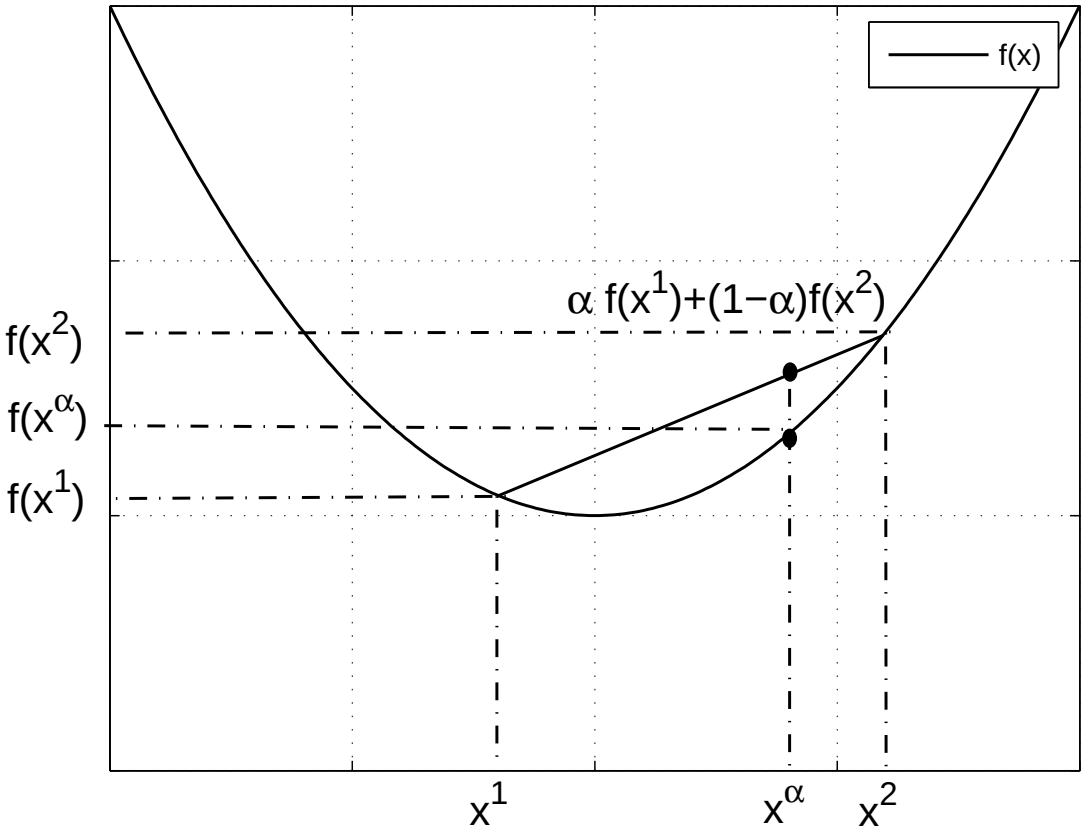

Figure: Illustration of the definition of a convex function in the one-dimensional case. Any straight line connecting two points on the graph lies "above" the graph. On the plot, $x_\alpha = \alpha x^1 + (1-\alpha)x^2$

Convex functions

- A function f is called concave if the negative of f is convex.
- A function is concave if it satisfies the property: For a given $\alpha \in [0, 1]$ and all $x^1 \in \mathbb{R}^n$ and $x^2 \in \mathbb{R}^n$ in the function domain,

$$f(\alpha x^1 + (1 - \alpha)x^2) \geq \alpha f(x^1) + (1 - \alpha)f(x^2).$$

Convex functions

- If the domain D of a convex function is not the entire space $\mathbb{R}^n$, then the set D satisfies the property,

$$\alpha x^1 + (1 - \alpha) x^2 \in D \tag{6}$$

  where $x^1 \in D$, $x^2 \in D$, and $0 \leq \alpha \leq 1$.
- The sets which satisfy (6) are called convex sets. Thus, the domains of convex (and concave) functions should be convex sets. Geometrically, a set is convex if it contains the straight line connecting any two points belonging to the set.

Convex functions

We summarize several important properties of convex functions:

1. Not all convex functions are differentiable. If a convex function is two times continuously differentiable, then the corresponding Hessian defined in (4) is a positive semidefinite matrix.
2. All convex functions are continuous if considered in an open set.
3. The sublevel sets

$$L_c = \{x : f(x) \leq c\}, \tag{7}$$

where c is a constant, are convex sets if f is a convex function. The converse is not true in general.

Convex functions

4. The local minima of a convex function are global. If a convex function f is twice continuously differentiable, then the global minimum is obtained in the points solving the first-order condition

$$\nabla f(x) = 0.$$

5. A sum of convex functions is a convex function:

$$f(x) = f_1(x) + f_2(x) + \ldots + f_k(x)$$

is a convex function if f_i, $i = 1, \ldots, k$ are convex functions.

Convex functions

- A simple example of a convex function is the linear function,

$$f(x) = a'x, \; x \in \mathbb{R}^n$$

where $a \in \mathbb{R}^n$ is a vector of constants. In fact, the linear function is the only function which is both convex and concave.

- As a more involved example, consider the following function,

$$f(x) = \frac{1}{2}x'Cx, \; x \in \mathbb{R}^n \tag{8}$$

where $C = \{c_{ij}\}_{i,j=1}^n$ is a $n \times n$ symmetric matrix. In this case, C is the covariance matrix.

Convex functions

- The function defined in (8) is called a quadratic function because writing the definition in terms of the components of the argument X, we obtain

$$f(x) = \frac{1}{2}\sum_{i=1}^{n} c_{ii}x_i^2 + \sum_{i \neq j} c_{ij}x_i x_j$$

  which is a quadratic function of the components x_i, $i = 1, \ldots, n$.
- The function in (8) is convex if and only if the matrix C is positive semidefinite. In fact, in this case the matrix C equals the Hessian matrix, $C = H$. Since the matrix C contains all parameters, we say that the quadratic function is defined by the matrix C.

Convex functions

Figures below illustrate the surface and contour lines of a convex and non-convex two-dimensional quadratic functions.

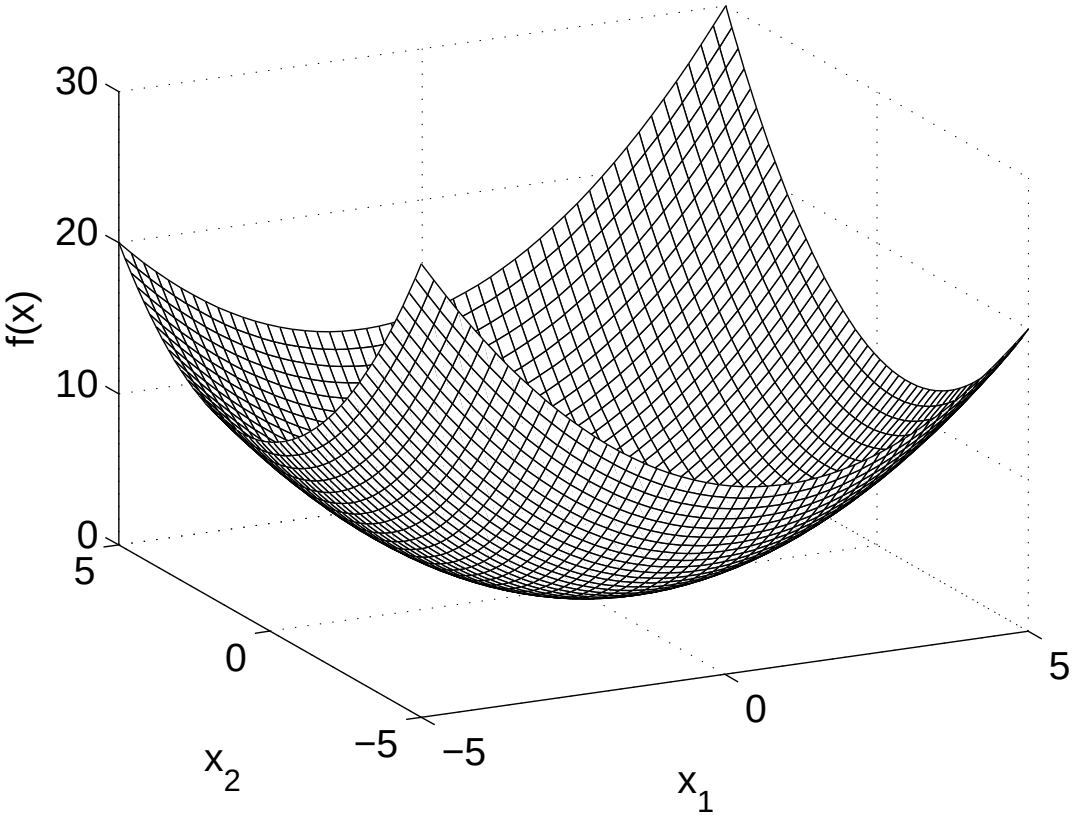

Figure: The surface of a two-dimensional convex quadratic function $f(x) = \frac{1}{2}x'Cx$

Convex functions

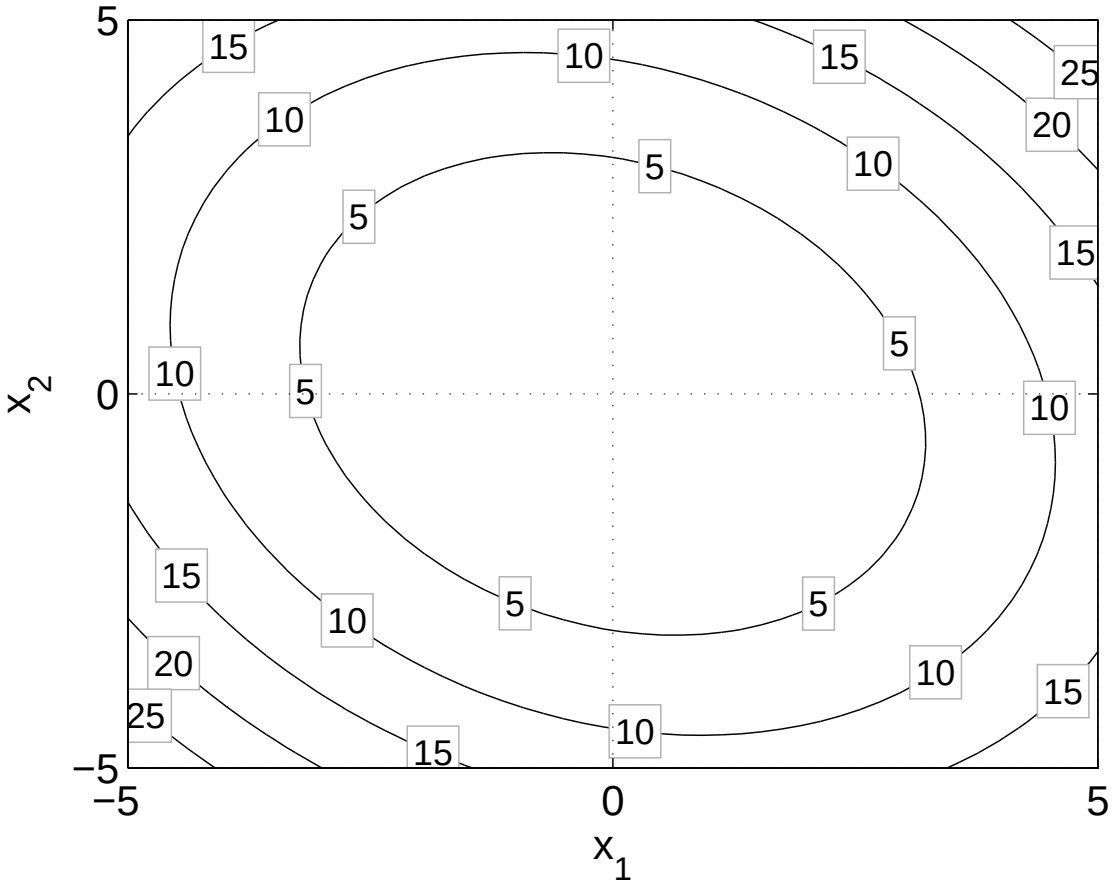

Figure: The contour lines of the two-dimensional convex quadratic function $f(x) = \frac{1}{2}x'Cx$.

Convex functions

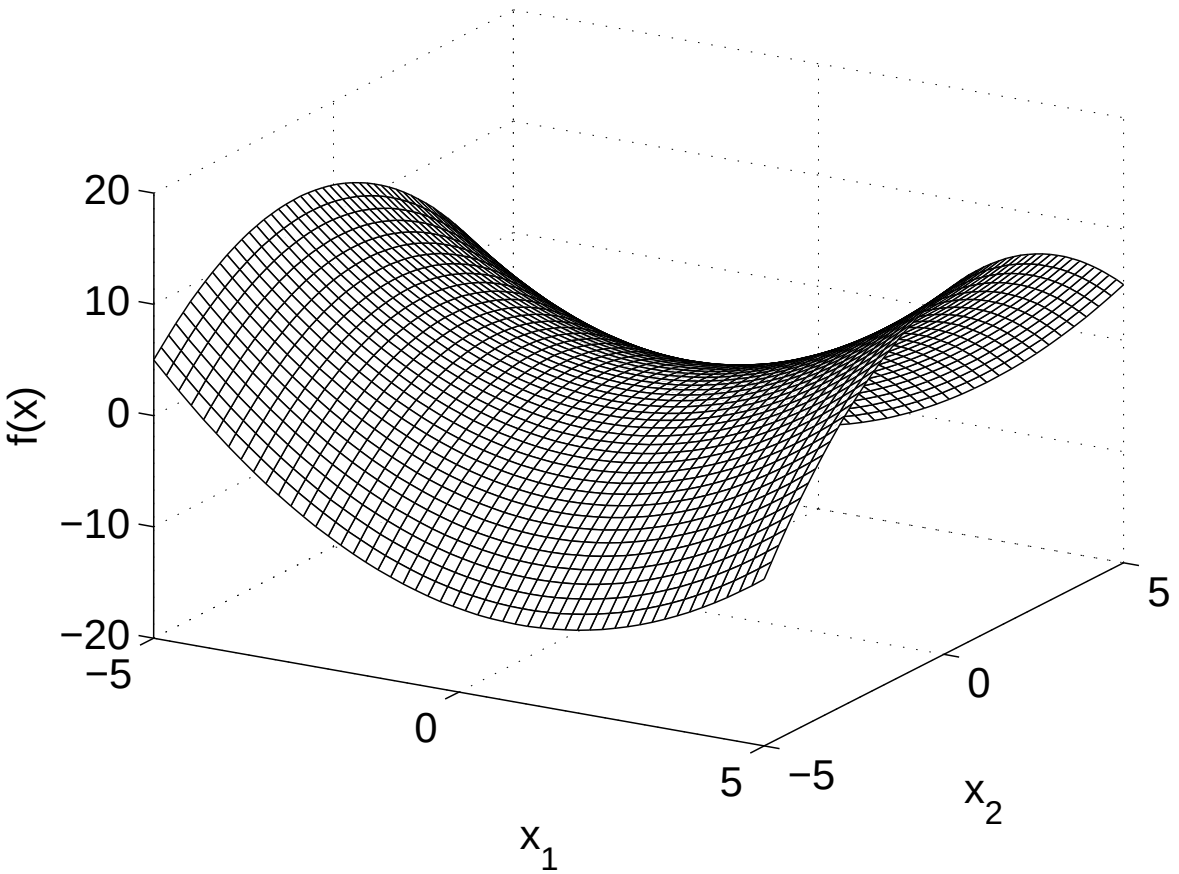

Figure: The surface of a non-convex two-dimensional quadratic function $f(x) = \frac{1}{2}x'Cx$. The point $(x_1, x_2) = (0, 0)$ is a saddle point.

Convex functions

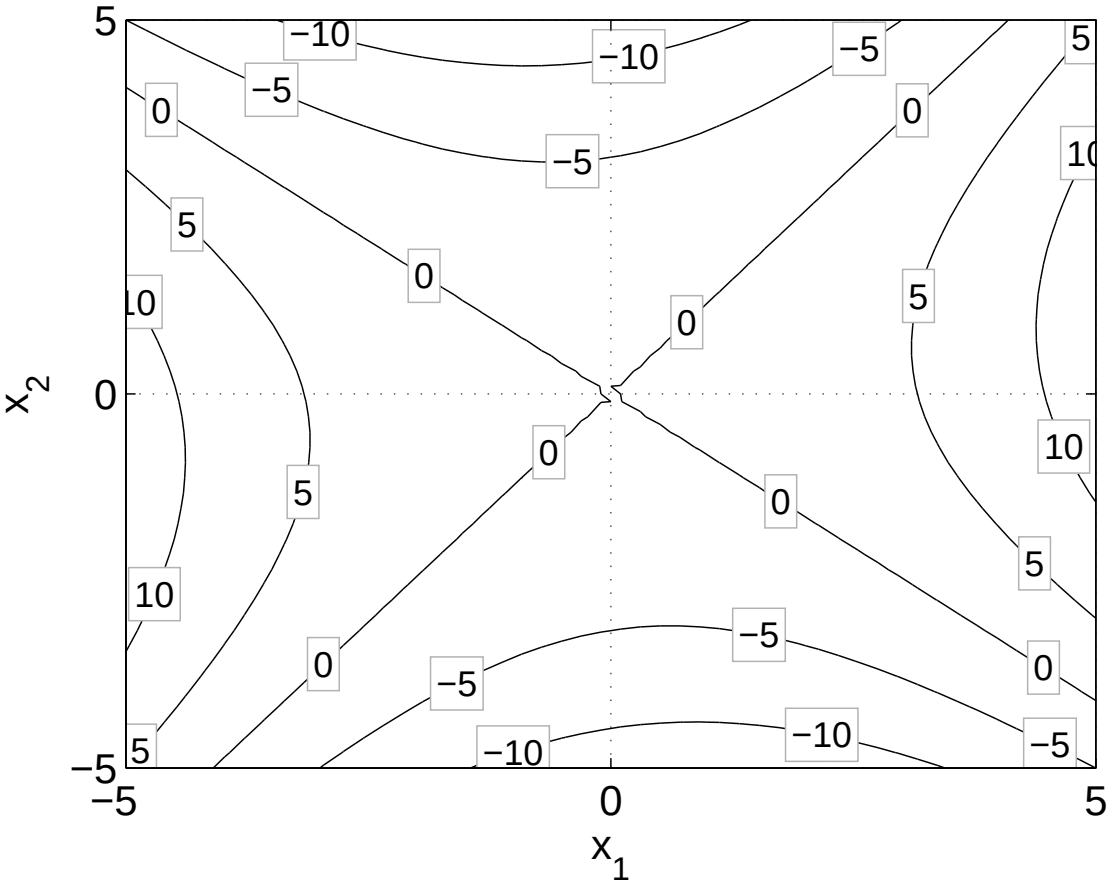

Figure: The contour lines of a non-convex two-dimensional quadratic function $f(x) = \frac{1}{2}x'Cx$. The point $(x_1, x_2) = (0, 0)$ is a saddle point.

Convex functions

- The contour lines of the convex function are concentric ellipses and a sublevel set L_c is represented by the points inside some ellipse.
- The convex quadratic function is defined by the matrix,

$$C = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}$$

and the non-convex quadratic function is defined by the matrix,

$$C = \begin{pmatrix} -1 & 0.4 \\ 0.4 & 1 \end{pmatrix}.$$

Convex functions

- A property of convex functions is that the sum of convex functions is a convex function. As a result of the preceding analysis, the function

$$f(x) = \lambda x' C x - a' x, \tag{9}$$

where $\lambda > 0$ and C is a positive semidefinite matrix, is a convex function as a sum of two convex functions.

Convex functions

- Let us use the properties of convex functions in order to solve the unconstrained problem of minimizing the function in (9),

$$\min_{x \in \mathbb{R}^n} \lambda x' C x - a' x$$

- This function is differentiable and we can search for the global minimum by solving the first-order conditions,

$$\nabla f(x) = 2\lambda C x - \mu = 0.$$

- Therefore, the value of x minimizing the objective function equals

$$x^0 = \frac{1}{2\lambda} C^{-1} \mu,$$

where C^{-1} denotes the inverse of the matrix C.

Quasi-convex functions

- Besides convex functions, there are other classes of functions with convenient optimal properties. An example of such a class is the class of quasi-convex functions.
- Formally, a function is called quasi-convex if all sublevel sets defined in (7) are convex sets. Alternatively, a function $f(x)$ is called quasi-convex if,

$$f(x^1) \geq f(x^2) \quad \text{implies} \quad f(\alpha x^1 + (1-\alpha)x^2) \leq f(x^1)$$

  where x^1 and x^2 belong to the function domain, which should be a convex set, and $0 \leq \alpha \leq 1$.
- A function f is called quasi-concave if $-f$ is quasi-convex.

Quasi-convex functions

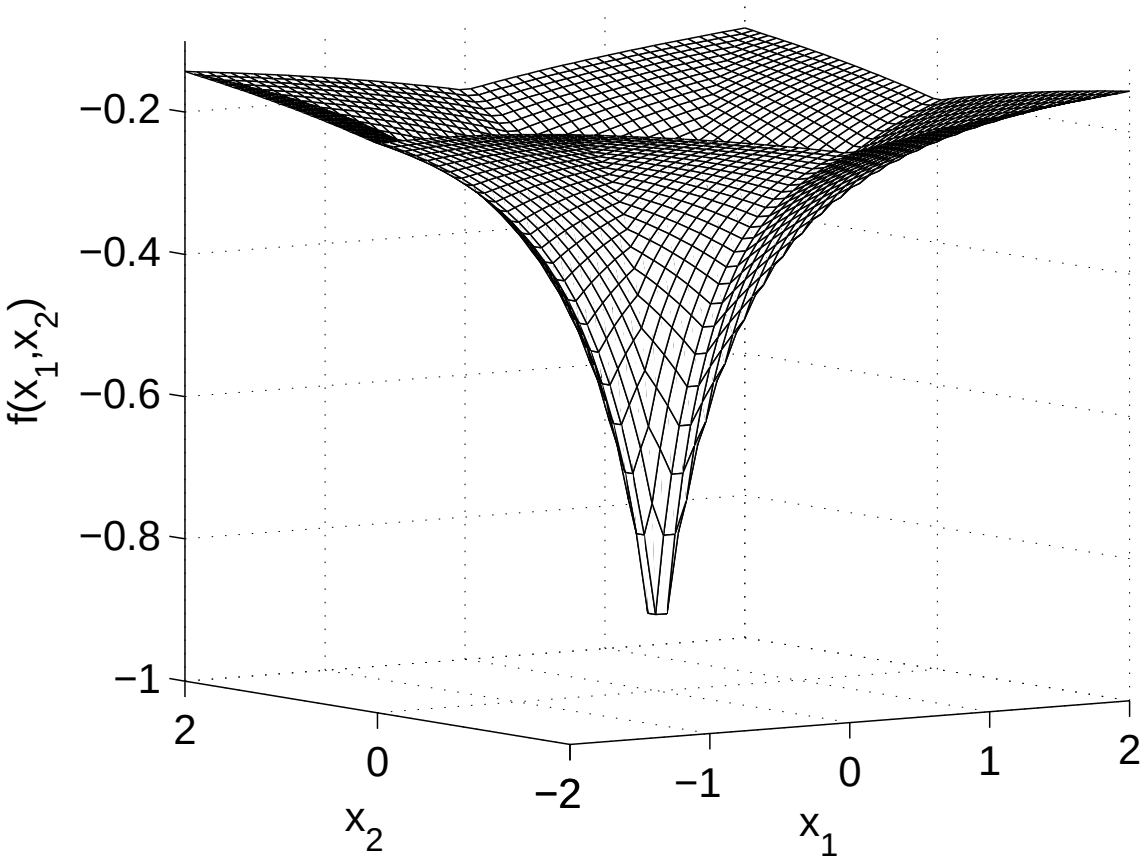

Figure: An example of a two-dimensional quasi-convex function $f(x_1, x_2)$. Even though the sublevel sets are convex, $f(x_1, x_2)$ is not a convex function.

Quasi-convex functions

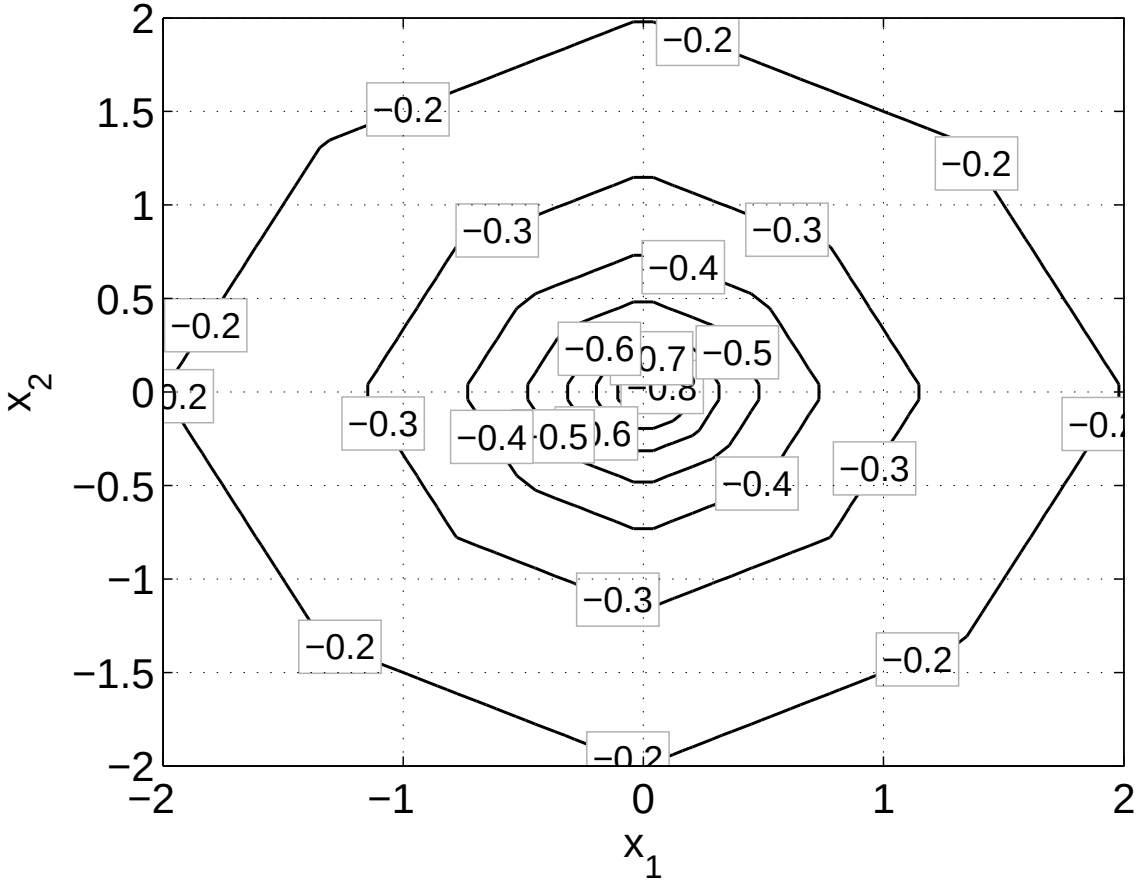

Figure: The contour lines of a two-dimensional quasi-convex function $f(x_1, x_2)$. Even though the sublevel sets are convex, $f(x_1, x_2)$ is not a convex function.

Quasi-convex functions

- A sublevel set is represented by all points inside some contour line.
- From a geometric viewpoint, the sublevel sets corresponding to the plotted contour lines are convex because any of them contains the straight line connecting any two points belonging to the set.
- Nevertheless, the function is not convex which becomes evident from the surface on the first plot. It is not guaranteed that a straight line connecting any two points on the surface will remain “above” the surface.

Quasi-convex functions

Several properties of the quasi-convex functions:

1. Any convex function is also quasi-convex. The converse is not true, which is demonstrated in figures on the slides 35,36.
2. In contrast to the differentiable convex functions, the first-order condition is not necessary and sufficient for optimality in the case of differentiable quasi-convex functions.
3. It is possible to find a sequence of convex optimization problems yielding the global minimum of a quasi-convex function. Its main idea is to find the smallest value of c for which the corresponding sublevel set L_c is non-empty. The minimal value of c is the global minimum which is attained in the points belonging to the sublevel set L_c.
4. Suppose that $g(x) > 0$ is a concave function and $f(x) > 0$ is a convex function. Then the ratio $g(x)/f(x)$ is a quasi-concave function and the ratio $f(x)/g(x)$ is a quasi-convex function.

Constrained optimization

- Solving practical issues, it is very often the case of imposing certain constraints for the optimal solution. For example, long-only portfolio optimization problems require that the portfolio weights, which represent the variables in optimization, should be non-negative and should sum up to one.
- This corresponds to a problem of the type,

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x'e = 1 \\ & x \geq 0, \end{aligned} \tag{10}$$

where

$f(x)$	is the objective function
$e \in \mathbb{R}^n$	is a vector of ones, $e = (1, \ldots, 1)$
$x'e$	equals the sum of all components of x, $x'e = \sum_i^n x_i$
$x \geq 0$	means that all components of the vector $x \in \mathbb{R}^n$ are non-negative

Constrained optimization

- In problem (10), we are searching for the minimum of the objective function by varying x only in the set

$$\mathbf{X} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} x'e = 1 \\ x \geq 0 \end{array} \right\}, \tag{11}$$

  which is also called the set of feasible points or the constraint set.
- A more compact notation, similar to the notation in the unconstrained problems, is sometimes used,

$$\min_{x \in \mathbf{X}} f(x)$$

  where $\mathbf{X}$ is defined in (11).

Constrained optimization

- We distinguish between different types of optimization problems depending on the assumed properties for the objective function and the constraint set.
- If the constraint set contains only equalities, the problem is easier to handle analytically. In this case, the method of Lagrange multipliers is applied.
- For more general constraint sets, when they are formed by both equalities and inequalities, the method of Lagrange multipliers is generalized by the Karush-Kuhn-Tucker conditions (KKT conditions).

Constrained optimization

- Like the first-order conditions we considered in unconstrained optimization problems, none of the two approaches lead to necessary and sufficient conditions for constrained optimization problems without further assumptions.
- One of the most general frameworks in which the KKT conditions are necessary and sufficient is that of convex programming. We have a convex programing problem if the objective function is a convex function and the set of feasible points is a convex set.
- As important sub-cases of convex optimization, linear programming and convex quadratic programming problems are considered.

Constrained optimization

In this part, we describe the following applications of constrained optimization:

- The method of Lagrange multipliers which is often applied to special types of mean-variance optimization problems in order to obtain closed-form solutions.
- The convex programming which is the framework for reward-risk analysis.

Lagrange multipliers

- Consider the following optimization problem in which the set of feasible points is defined by a number of equality constraints,

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & h_1(x) = 0 \\ & h_2(x) = 0 \\ & \dots \\ & h_k(x) = 0. \end{aligned} \tag{12}$$

- The functions $h_i(x)$, $i = 1, \dots, k$ build up the constraint set.

Remark: Even though the right hand-side of the equality constraints is zero in the classical formulation of the problem given in (12), this is not restrictive. If in a practical problem the right hand-side happens to be different than zero, it can be equivalently transformed, for example:

$$\{x \in \mathbb{R}^n : v(x) = c\} \qquad \Longleftrightarrow \qquad \{x \in \mathbb{R}^n : h_1(x) = v(x) - c = 0\}.$$

Lagrange multipliers

- In order to illustrate the necessary condition for optimality valid for (12), let us consider the following two-dimensional example:

$$\begin{array}{ll} min_{x \in \mathbb{R}^2} & \frac{1}{2} x' C x \\ \text{subject to} & x' e = 1. \end{array} \tag{13}$$

where the matrix is,

$$C = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}.$$

Lagrange multipliers

- The objective function is a quadratic function and the constraint set contains one linear equality.
- The surface of the objective function and the constraint are shown on the plot on the *nexr slide*.
- The black line on the surface shows the function values of the feasible points. Geometrically, solving problem reduces to finding the lowest point of the black curve on the surface.

Lagrange multipliers

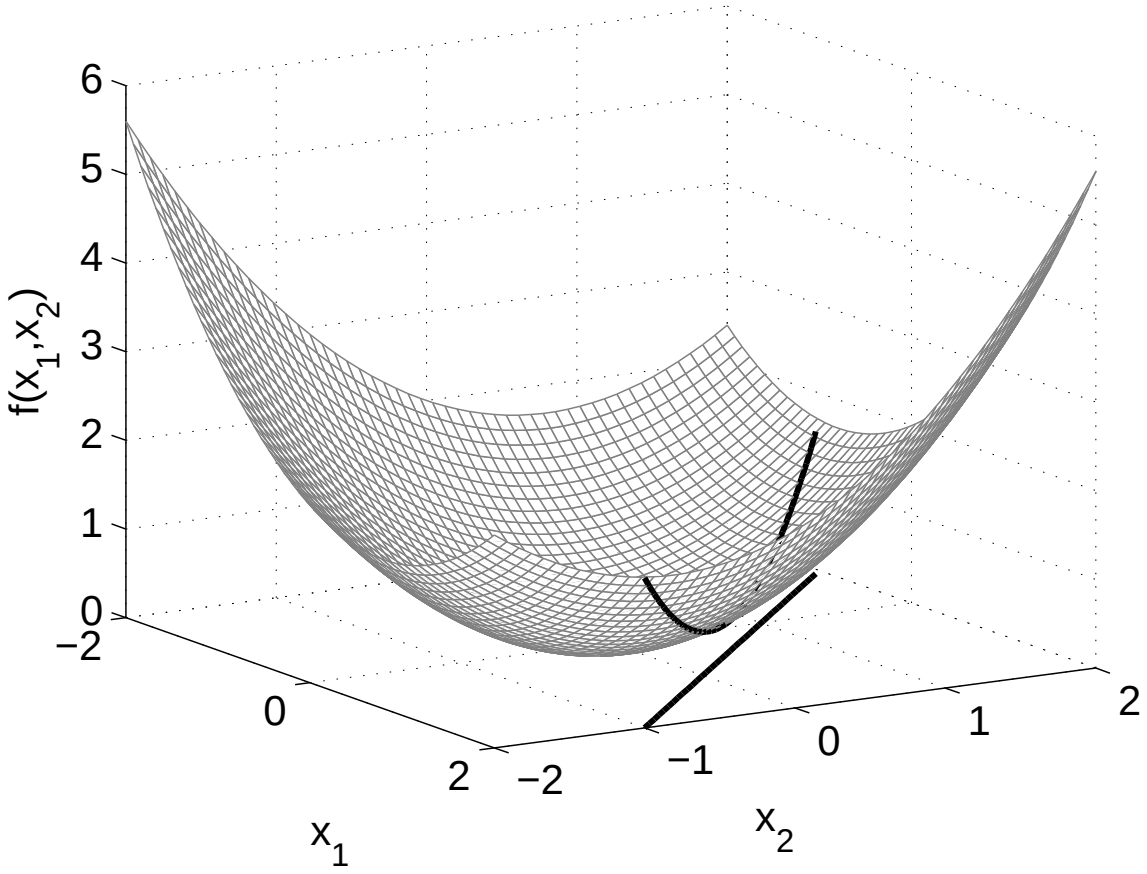

Figure: The plot shows the surface of a two-dimensional quadratic objective function and the linear constraint $x_1 + x_2 = 1$. The black curve on the surface shows the objective function values of the points satisfying the constraint.

Lagrange multipliers

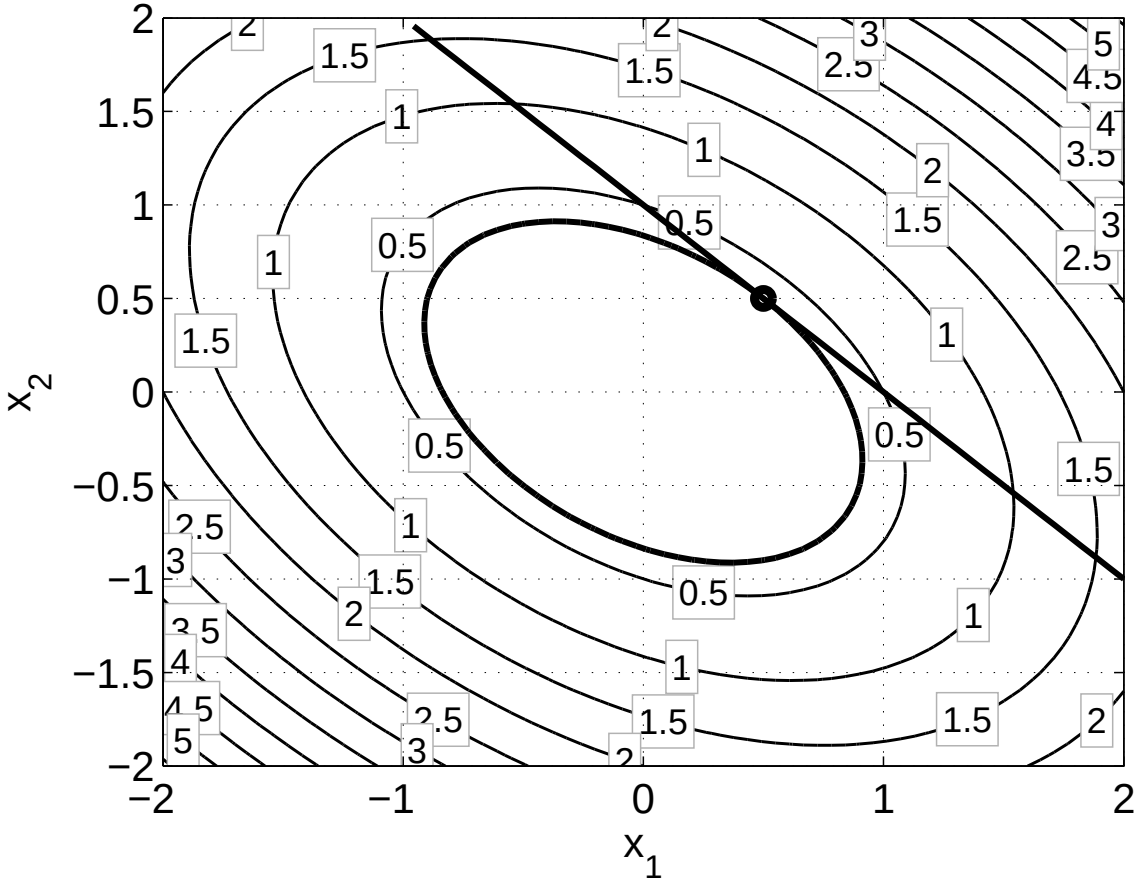

Figure: The plot shows the tangential contour line to the constraint $x_1 + x_2 = 1$. The black dot indicates the position of the point in which the objective function attains its minimum subject to the constraints.

Lagrange multipliers

- The fact the minimum is attained where a contour line is tangential to the curve defined by the non-linear equality constraints is expressed in the following way:

  The gradient of the objective function at the point yielding the minimum is proportional to a linear combination of the gradients of the functions defining the constraint set. Formally, this is stated as,

$$\nabla f(x^0) - \mu_1 \nabla h_1(x^0) - \ldots - \mu_k \nabla h_k(x^0) = 0. \tag{14}$$

  where μ_i, $i = 1, \ldots, k$ are some real numbers called Lagrange multipliers and the point x^0 is such that $f(x^0) \leq f(x)$ for all x which are feasible.

Lagrange multipliers

- Note that if there are no constraints in the problem, then (14) reduces to the first-order condition. Therefore, the system of equations behind (14) can be viewed as a generalization of the first-order condition in the unconstrained case.
- The method of Lagrange multipliers basically associates a function to the problem in (12) such that the first-order condition for unconstrained optimization for that function coincides with (14).

Lagrange multipliers

The method of Lagrange multiplier consists of the following steps.

1. Given the problem in (12), construct the following function

$$L(x, \mu) = f(x) - \mu_1 h_1(x) - \ldots - \mu_k h_k(x) \tag{15}$$

   where $\mu = (\mu_1, \ldots, \mu_k)$ is the vector of Lagrange multipliers. The function $L(x, \mu)$ is called the Lagrangian corresponding to problem (12).

Lagrange multipliers

2. Calculate the partial derivatives with respect to all components of x and μ and set them equal to zero,

$$\begin{aligned} \frac{\partial L(x,\mu)}{\partial x_i} &= \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^{k} \mu_j \frac{\partial h_j(x)}{\partial x_i} = 0, \quad i = 1, \ldots, n \\ \frac{\partial L(x,\mu)}{\partial \mu_m} &= h_m(x) = 0, \quad m = 1, \ldots, k \end{aligned} \tag{16}$$

Basically, the system of equations (16) corresponds to the first-order conditions for unconstrained optimization written for the Lagrangian as a function of both x and μ, $L : \mathbb{R}^{n+k} \to \mathbb{R}$.

Lagrange multipliers

3. Solve the system of equalities in (16) for x and μ.

   Note that even though we are solving the first-order condition for unconstrained optimization of $L(x, \mu)$, the solution (x^0, μ^0) of (16) is not a point of local minimum or maximum of the Lagrangian. In fact, the solution (x^0, μ^0) is a saddle point of the Lagrangian.

Lagrange multipliers

- The first n equations in (16) make sure that the relationship between the gradients given in (14) is satisfied.
- The following k equations in (16) make sure that the points are feasible.
- As a result, all vectors x solving (16) are feasible and the gradient condition is satisfied in them. Therefore, the points which solve the optimization problem (12) are among the solutions of the system of equations given in (16).

Lagrange multipliers

- This analysis suggests that the method of Lagrange multipliers provides a necessary condition for optimality.
- Under certain assumptions for the objective function and the functions building up the constraint set, (16) turns out to be a necessary and sufficient condition.

Example

If $f(x)$ is a convex and differentiable function and $h_i(x)$, $i = 1, \ldots, k$ are affine functions, then the method of Lagrange multipliers identifies the points solving (12).

Figures on the slide 47,48 illustrate a convex quadratic function subject to a linear constraint. In this case, the solution point is unique.

Convex programming

The general form of convex programming problems is the following,

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_j(x) = 0, \quad j = 1, \ldots, k \end{aligned} \tag{17}$$

where

$f(x)$	is a convex objective function
$g_1(x), \ldots, g_m(x)$	are convex functions defining the inequality constraints
$h_1(x), \ldots, h_k(x)$	are affine functions defining the equality constraints

Convex programming

- Generally, without the assumptions of convexity, problem (17) is more involved than (12) because besides the equality constraints, there are inequality constraints.
- The KKT condition, generalizing the method of Lagrange multipliers, is only a necessary condition for optimality in this case. However, adding the assumption of convexity makes the KKT condition necessary and sufficient.
- Note that, similar to problem (12), the fact that the right hand-side of all constraints is zero is non-restrictive. The limits can be arbitrary real numbers.

Convex programming

- Consider the following two-dimensional optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & \tfrac{1}{2} x' C x \\ \text{subject to} \quad & (x_1 + 2)^2 + (x_2 + 2)^2 \leq 3 \end{aligned} \tag{18}$$

in which

$$C = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}.$$

Convex programming

- The objective function is a two-dimensional convex quadratic function and the function in the constraint set is also a convex quadratic function.
- In fact, the boundary of the feasible set is a circle with a radius of $\sqrt{3}$ centered at the point with coordinates $(-2, -2)$.
- The plots on the *next slides* show the surface of the objective function, its contour lines and the set of feasible points.

Convex programming

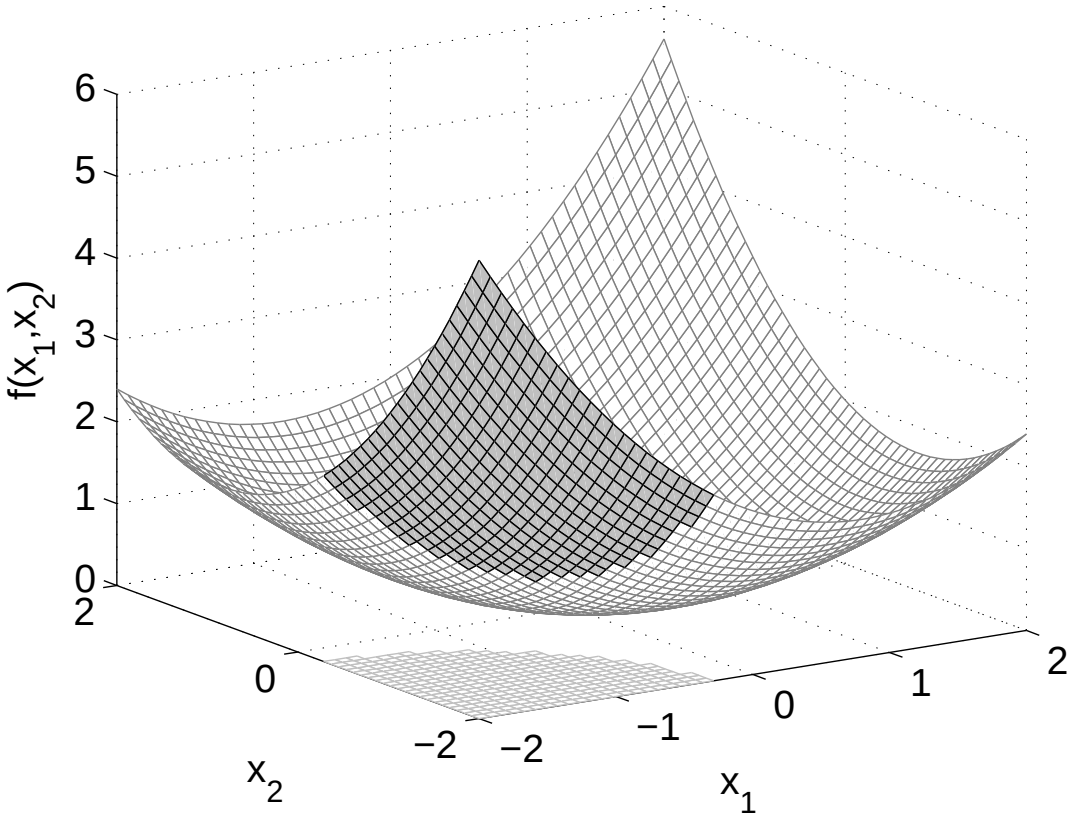

Figure: The plot shows the surface of a two-dimensional convex quadratic function and a convex quadratic constraint. The shaded section on the surface corresponds to the feasible points. Solving problem (18) reduces to finding the lowest point on the shaded part of the surface.

Convex programming

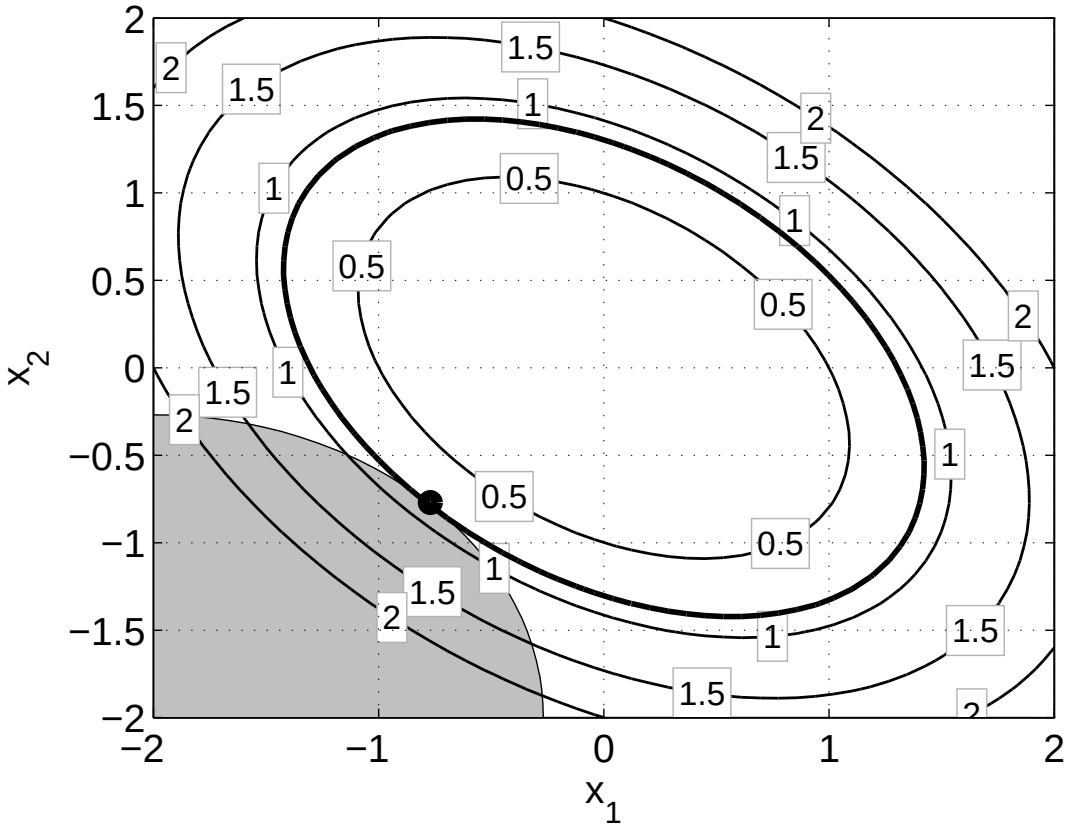

Figure: The plot shows the tangential contour line to the feasible set, which is in gray. Geometrically, the point in the feasible set yielding the minimum of the objective function is positioned where a contour line only touches the constraint set. The position of this point is marked with a black dot and the tangential contour line is given in bold.

Convex programming

- Note that the solution points of problems of the type (18) can happen to be not on the boundary of the feasible set but in the interior.
- For example, suppose that the radius of the circle defining the boundary of the feasible set in (18) is a larger number such that the point $(0, 0)$ is inside the feasible set.
- Then, the point $(0, 0)$ is the solution to problem (18) because at this point the objective function attains its global minimum.

Convex programming

- In the two-dimensional case, when we can visualize the optimization problem, geometric reasoning guides us to finding the optimal solution point.
- In a higher dimensional space, plots cannot be produced and we rely on the analytic method behind the KKT conditions.
- The KKT conditions corresponding to the convex programming problem (17) are the following:

$$
\begin{aligned}
&\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{j=1}^{k} \mu_j \nabla h_j(x) = 0 \\
&g_i(x) \leq 0 \quad i = 1, \ldots, m \\
&h_j(x) = 0 \quad j = 1, \ldots, k \\
&\lambda_i g_i(x) = 0, \quad i = 1, \ldots, m \\
&\lambda_i \geq 0, \quad i = 1, \ldots, m.
\end{aligned}
\tag{19}
$$

Convex programming

- A point x^0 such that (x^0, λ^0, μ^0) satisfies (19) is the solution to problem (17).
- Note that if there are no inequality constraints, then the KKT conditions reduce to (16) in the method of Lagrange multipliers. Therefore, the KKT conditions generalize the method of Lagrange multipliers.

Convex programming

- The gradient condition in (19) has the same interpretation as the gradient condition in the method of Lagrange multipliers. The set of constraints,

$$\begin{aligned} g_i(x) &\leq 0 \quad i = 1, \ldots, m \\ h_j(x) &= 0 \quad j = 1, \ldots, k \end{aligned}$$

guarantee that a point satisfying (19) is feasible.

Convex programming

- The next conditions

$$\lambda_i g_i(x) = 0, \quad i = 1, \ldots, m$$

  are called complementary slackness conditions.
- If an inequality constrain is satisfied as a strict inequality, then the corresponding multiplier λ_i turns into zero according to the complementary slackness conditions.
- Then the corresponding gradient $\nabla g_i(x)$ has no significance in the gradient condition. This reflects the fact that the gradient condition concerns only the constraints satisfied as equalities at the solution point.

Linear programming

- Optimization problems are said to be linear programming problems if the objective function is a linear function and the feasible set is defined by linear equalities and inequalities.
- Since all functions are linear, they are also convex which means that linear programming problems are also convex problems.
- The definition of linear programming problems in standard form is the following:

$$\begin{aligned} \min_{x} \quad & c'x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0, \end{aligned} \tag{20}$$

where A is a $m \times n$ matrix of coefficients, $c = (c_1, \ldots, c_n)$ is a vector of objective function coefficients, and $b = (b_1, \ldots, b_m)$ is a vector of real numbers. As a result, the constraint set contains m inequalities defined by linear functions.

Linear programming

- The feasible points defined by means of linear equalities and inequalities are also said to form a polyhedral set.
- Figure on the *next slide* shows an example of a two-dimensional linear programming problem which is not in standard form as the two variables may become negative.

Linear programming

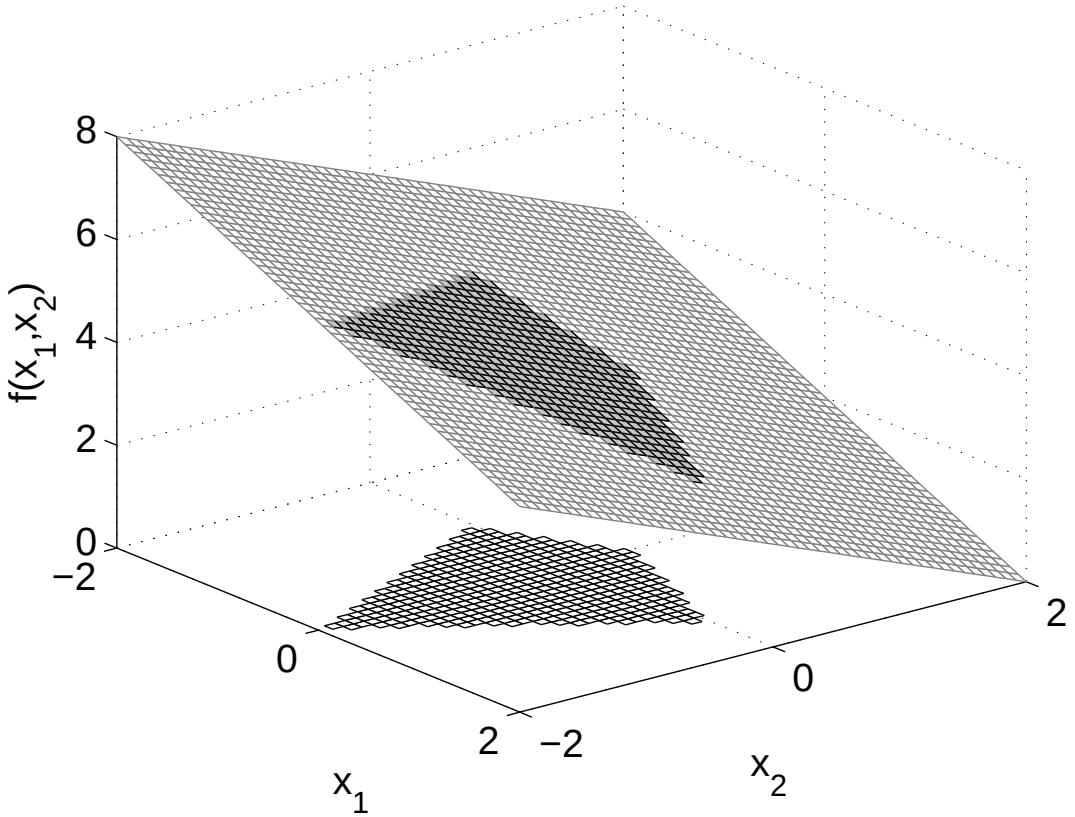

Figure: The plot shows the surface of a linear function and a polyhedral feasible set. The shaded section on the surface corresponds to the feasible points. Solving problem (20) reduces to finding the lowest point in the shaded area on the surface.

Linear programming

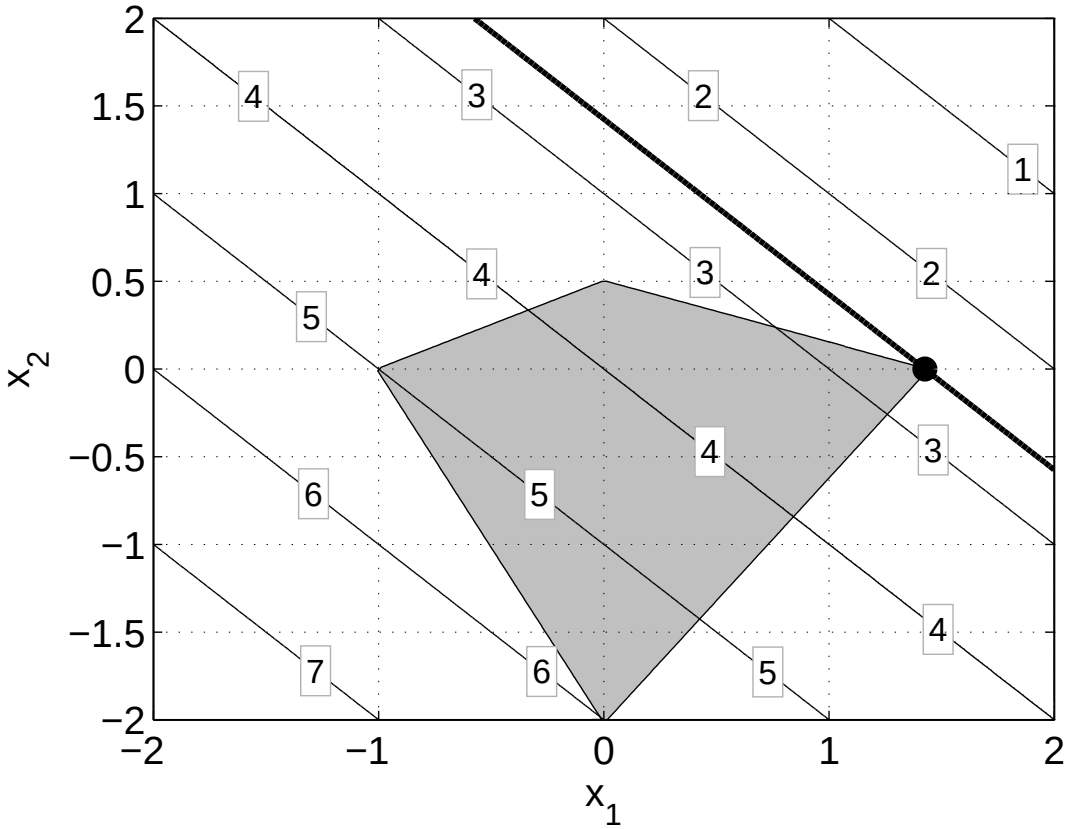

Figure: The plot shows the tangential contour line to the feasible set. The contour lines are parallel straight lines because the objective function is linear. The point in which the objective function attains its minimum is marked with a black dot.

Linear programming

- A general result in linear programming is that, on condition that the problem is bounded, the solution is always at the boundary of the feasible set and, more precisely, at a vertex of the polyhedron.
- Problem (20) may become unbounded if the polyhedral set is unbounded and there are feasible points such the objective function can decrease indefinitely.
- We can summarize that, generally, due to the simple structure of (20), there are three possibilities:
  1. The problem is not feasible, because the polyhedral set is empty
  2. The problem is unbounded
  3. The problem has a solution at a vertex of the polyhedral set

Linear programming

- The polyhedral set has a finite number of verices and an algorithm can be devised with the goal of finding a vertex solving the optimization problem in a finite number of steps.
- This is the basic idea behind the simplex method which is an efficient numerical approach to solving linear programming problems. Besides the simplex algorithm, there are other more contemporary methods such as the interior point method for example.

Application of linear programming:

- A few classes of practical problems which are solved by the method of linear programming include the *transportation problem*, the *transshipment problem*, the *network flow problem* and so on.[1].

[1] Dantzig (1998) provides an excellent background on the theory and application of linear programing

Quadratic programming

Another class of problems is quadratic programming problems.
It contains optimization problems with a quadratic objective function and linear equalities and inequalities in the constraint set,

$$\begin{aligned} \min_{x} \quad & c'x + \tfrac{1}{2}x'Hx \\ \text{subject to} \quad & Ax \leq b, \end{aligned} \tag{21}$$

where

$c = (c_1, \ldots, c_n)$	is a vector of coefficients defining the linear part of the objective function
$H = \{h_{ij}\}_{i,j=1}^{n}$	is a $n \times n$ matrix defining the quadratic part of the objective
$A = \{a_{ij}\}$	is a $k \times n$ matrix defining k linear inequalities in the constraint set
$b = (b_1, \ldots, b_k)$	is a vector of real numbers defining the right hand-side of the linear inequalities

Quadratic programming

- In optimal portfolio theory, mean-variance optimization problems in which portfolio variance is in the objective function are quadratic programming problems.
- From the point of view of optimization theory, problem (21) is a convex optimization problem if the matrix defining the quadratic part of the objective function is positive semidefinite. In this case, the KKT conditions can be applied to solve it.

Svetlozar T. Rachev, Stoyan Stoyanov, and Frank J. Fabozzi
Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures
John Wiley, Finance, 2007.

Chapter 2.