Continuous Optimization

Chapter 1: Basics of Optimization

1 Optimization problems

In this course we will focus on finite-dimensional continuous optimization problems

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \leq 0 \quad i = 1, \ldots, m \\
& h_j(x) = 0 \quad j = 1, \ldots, p
\end{aligned}
\tag{1}
$$

- $x \in \mathbb{R}^n$: vector of variables to be chosen (n scalar variables $x_1, \ldots, x_n$)

- f: objective function to be minimized

- $g_1, \ldots, g_m$: inequality constraint functions to be satisfied

- $h_1, \ldots, h_p$: equality constraint functions to be satisfied

The feasible set can also be denoted by $S = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \ i = 1, \ldots, m, \ \ h_j(x) = 0 \ j = 1, \ldots, p\}$.

1.1 Taxonomy

Concerning the dimensionality, optimization problems can be

- infinite-dimensional: when the space of variables is infinite dimensional, e.g., elements can be vectors of functions of one ore more variables;

- finite-dimensional: when the space of variables is finite-dimensional, e.g., $\mathbb{R}^n$.

Concerning the granularity of the space, optimization problems can be

- discrete: when at least one variable is an integer;

- continuous: when all variables involved are continuous.

Concerning the hypothesis of linearity, optimization problems are also called

- linear programming: when the objective function is linear and the feasible set is defined by a system of linear equalities and inequalities;

- nonlinear problems: when the objective function is nonlinear and/or at least one constraint is defined by a nonlinear function.

Concerning the hypothesis of differentiability, we can distinguish between

- smooth optimization: when the functions defining the objective and the constraints are continuously differentiable;

- nonsmooth optimization problems: when the functions defining the objective and the constraints are continuous, but typically not differentiable in all points of the feasible set. Weaker notions of derivatives are used to solve this problem.

- derivative-free optimization: this branch of optimization consider cases in which evaluating the objective function/constraints is very costly, in particular the functions defining the objective and the constraints might not be continuous. A similar field is that of black-box optimization, in which the optimizer assumes to not have any information on the objective function or on the constraints.

Concerning the hypothesis of convexity, we can distinguish between

- convex optimization, when the functions defining the objective and the constraints are all convex;

- nonconvex optimization, when the functions defining the objective and the constraints might be nonconvex.

Concerning the uncertainties in the model, we can distinguish between

- stochastic optimization, when the uncertainties (i.e., random variables) are maintained in the model and their quantification (e.g., expectation, variance) is exploited in optimization algorithm;

- deterministic optimization, when there are no uncertainties, or we assume that this is the case.

Example 1.1 (The transportation problem (Linear Programming)). *Consider a company producing and selling a certain commodity. The company has a set of N sources, where the commodity is produced and a set of M demand centers where the commodity is sold. At each source $i \in \{1, \dots, N\}$, a given quantity q_i of commodity is available. Each demand center $j \in \{1, \dots, M\}$ requires a given quantity d_j. We indicate by c_{ij} the cost per unit of transporting the commodity from source i to demand center j.*

The problem is to determine the quantity to be transported from each source to each demand center in such a way that

- *the availability constraints at the sources are satisfied;*

- *the requirements of the demand centers are satisfied;*

- *the total transportation cost is minimized.*

To formulate the problem, we indicate by x_{ij} the quantity of the commodity transported from source i to demand center j. Then the optimization problem is

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{N} \sum_{j=1}^{M} c_{ij} x_{ij} \\
s.t. \quad & \sum_{j=1}^{M} x_{ij} \leq q_i \quad i = 1, \dots, N \\
s.t. \quad & \sum_{j=1}^{M} x_{ij} = d_j \quad j = 1, \dots, M \\
& x_{ij} \geq 0 \qquad i = 1, \dots, N \ \ j = 1, \dots, M.
\end{aligned}
$$

Example 1.2 (Training of Neural Networks (Unconstrained Optimization)). *Let us start from describing a single (artificial) neuron. In particular, this unit takes as an input a vector and performs a scalar product with the weights of the neuron. In a second step, this value passes through a thresholding function σ called activation function. Inspired by real neurons, also this unit propagates its output only if the weighted sum of its inputs (in the natural neuron, the potential difference) is over a certain threshold. Neural networks stack various neurons in parallel to build a layer and, in turn, concatenate various layer (deep networks). This model is very expressive, in fact already a wide-enough 2-layer neural network is able to approximate any possible function of the input.*

Neural networks are the most widely used machine learning model and they can be formally defined as follows.

- $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{M}$ *be a training set of M instances*
  - $x_i \in \mathbb{R}^d$ *d is the amount of features of the dataset (e.g., pixels)*
  - $y_i \in \mathbb{R}$

- *the j-th layer applied on a generic input n_{j-1}-dimensional input x:*
  $g_j(x) := \sigma(W_j x + b_j)$, $W_j \in \mathbb{R}^{n_{j-1} \times n_j}$, $b_j \in \mathbb{R}^{n_j}$, σ *is applied component-wise,*
  e.g., ReLU $\sigma(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{otherwise} \end{cases}$

- *the weights:* $w = \begin{pmatrix} vec(W_1) \\ \vdots \\ vec(W_L) \end{pmatrix} \in \mathbb{R}^n$

- *the network applied on the i-th instance:* $h_i(w) := g_L \circ \cdots \circ g_1(x_i)$

- *the losses:* $f_i(w) := \mathcal{L}(h_i(w), y_i)$, *e.g.,* $\mathcal{L}(h_i(w), y_i) = (h_i(w) - y_i)^2$

$$\min_{w \in \mathbb{R}^n} f(w) = \frac{1}{M} \sum_{i=1}^{M} f_i(w).$$

Example 1.3 (Portfolio Optimization (Convex Quadratic Programming))**.** *Portfolio selection theory studies how to allocate an investor's available capital into a prefixed set of assets with the aims of maximizing the expected return and minimizing the investment risk. Let n be the number of available assets, let $\mu \in \mathbb{R}^n$ be the vector of expected returns of the assets, and let $Q \in \mathbb{R}^{n \times n}$ be a symmetric positive semidefinite matrix whose generic element q_{ij} is the covariance of returns of assets i and j. We assume that the available (normalized) capital is fully invested. Then, let $x \in \mathbb{R}^n$ be the vector of decision variables, where x_i is the fraction of the available capital to be invested into asset i, with $i = 1, \ldots, n$.*

By this notation, $\mu^T x$ is the expected return of the portfolio and $x^T Q x$ is the covariance of the portfolio which can be used as a measure of the risk connected with the investment (diversification of the portfolio). In the traditional Markowitz portfolio selection model [1], the optimization problem is stated as the following convex quadratic programming problem

$$\begin{aligned} \min \ & x^T Q x \\ s.t. \ & \mu^T x \geq \beta \\ & e^T x = 1 \\ & x \geq 0, \end{aligned}$$

where β is the desired expected return of the portfolio and $e \in \mathbb{R}^n$ denotes the column vector of all ones.

Example 1.4 (Shortest Path (Discrete Optimization, not covered by this course))**.** *Let $G = (V, E)$ be a graph with V the set of vertices and E the set of edges (directed, $(i, j) \neq (j, i)$, with i, j being two vertices) with $n = |E|$. Let $c_{ij} \geq 0$ be the cost of choosing the edge (i, j), e.g., travel time, travel fee, difference in hight, and let $x_{ij} \in \{0, 1\}^n$ be the variable telling if we select or not the edge (i, j). Also, let s be the starting node and d the destination, then the optimization problem can be stated as*

$$\begin{aligned} \min \ & c^T x \\ s.t. \ & x \geq 0 \end{aligned}$$

$$\sum_j x_{ij} - \sum_j x_{ji} = \begin{cases} 1 & i = s \\ 0 & i \neq s, d \\ -1 & i = d \end{cases} \quad \forall i \in V.$$

2 Mathematical Preliminaries

2.1 The space $\mathbb{R}^n$

The vector space $\mathbb{R}^n$ is the set of n-dimensional column vectors with real components endowed with the component-wise addition operator

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

and the scalar product

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

The standard basis of $\mathbb{R}^n$ is $e_1, e_2, \ldots, e_n$, where e_i is the n-length column vector whose ith component is 1 while all the others are 0.

The nonnegative orthant $\mathbb{R}^n_+ := \{(x_1, x_2, \ldots, x_n)^T : x_1, x_2, \ldots, x_n \geq 0\}$.

The positive orthant $\mathbb{R}^n_{++} := \{(x_1, x_2, \ldots, x_n)^T : x_1, x_2, \ldots, x_n > 0\}$.

The set of all real-valued matrices of order $m \times n$ is denoted by $\mathbb{R}^{m \times n}$.

For given $x, y \in \mathbb{R}^n$, the closed line segment between x and y is a subset of $\mathbb{R}^n$ denoted by $[x, y]$ and defined as

$$[x, y] := \{x + a(y - x) : \alpha \in [0, 1]\}.$$

The open line segment (x, y) is similarly defined as

$$(x, y) := \{x + a(y - x) : \alpha \in (0, 1)\}.$$

The identity matrix will be denoted by Id, where its dimension will be clear from the context.

2.2 Inner Products

Definition 2.1 (Inner product). *An inner product on $\mathbb{R}^n$ is a map $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with the following properties*

1. *Symmetry: $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in \mathbb{R}^n$;*

2. *Additivity: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad \forall x, y, z \in \mathbb{R}^n$;*

3. *Homogeneity: $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle \quad \forall x, y \in \mathbb{R}^n$;*

4. *Positive definiteness: $\langle x, x \rangle \geq 0 \quad \forall x \in \mathbb{R}^n$ and $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.*

In this course the inner product we will use is the standard dot product:

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i \quad \forall x, y \in \mathbb{R}^n$$

Notice that the dot product is not the only inner product in $\mathbb{R}^n$, for instance let $w \in \mathbb{R}^n_{++}$, it is easy to show that the following weighted dot product is also an inner product

$$\langle x, y \rangle_w = \sum_{i=1}^{n} w_i x_i y_i.$$

2.3 Vector norms

Definition 2.2 (Vector Norm). *A norm $||.||$ on $\mathbb{R}^n$ is a function $||.|| : \mathbb{R}^n \to \mathbb{R}$ satisfying the following:*

1. *Nonnegativity: $||x|| \geq 0 \quad \forall x \in \mathbb{R}^n$ and $||x|| = 0 \Leftrightarrow x = 0$;*

2. *Positive Homogeneity: $||\lambda x|| = |\lambda| ||x|| \quad \forall x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$;*

3. *Triangle Inequality: $||x + y|| \leq ||x|| + ||y|| \quad \forall x, y \in \mathbb{R}^n$.*

One natural way to generate a norm on $\mathbb{R}^n$ is to take any inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$ and define the associated norm

$$||x|| := \sqrt{\langle x, x \rangle} \quad \forall x \in \mathbb{R}^n$$

which can be proved to be a norm. If the inner product is the dot product, then the associated norm is the Euclidean norm or l_2 norm, i.e., $|| \cdot ||_2$. In the rest of the course, the default norm will be the Euclidean norm and the subscript 2 will be omitted. This norm belongs to the class of l_p norms (for $p \geq 1$, with $0 \leq p < 1$ these are quasi-norms) defined by

$$||x||_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}.$$

By computing the limit of the l_p-norm with $p \to \infty$ we achieve the l_∞ norm, i.e.,

$$||x||_\infty := \max_{i \in [n]} |x_i|.$$

where $[n] := \{1, \ldots, n\}$ for any $n \in \mathbb{N}$.

An important inequality connecting the dot product of two vectors and their norms is the Cauchy–Schwarz inequality, which will be used frequently throughout the book.

Lemma 2.1 (Cauchy–Schwarz inequality). *For any $x, y \in \mathbb{R}^n$,*

$$|x^T y| \leq ||x|| \cdot ||y||.$$

Equality is satisfied if and only if x and y are linearly dependent.

2.4 Matrix norms

Similarly to vector norms, we can define the concept of a matrix norm.

Definition 2.3 (Operator/Matrix Norm). *A norm $||.||$ on $\mathbb{R}^{m \times n}$ is a function $||.|| : \mathbb{R}^{m \times n} \to \mathbb{R}$ satisfying the following:*

1. *Nonnegativity: $||A|| \geq 0 \quad \forall A \in \mathbb{R}^{m \times n}$ and $||A|| = 0 \Leftrightarrow A = 0$;*

2. *Positive Homogeneity: $||\lambda A|| = |\lambda| ||A|| \quad \forall A \in \mathbb{R}^{m \times n}$ and $\lambda \in \mathbb{R}$;*

3. *Triangle Inequality: $||A + B|| \leq ||A|| + ||B|| \quad \forall A, B \in \mathbb{R}^{m \times n}$.*

Given a matrix $A \in \mathbb{R}^{m \times n}$ and two norms $||.||_p$ and $||.||_q$ on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, the induced operator/matrix norm $||A||_{a \to b}$ is defined by

$$||A||_{p \to q} = \max_{||x||_p = 1} ||Ax||_q.$$

It can be shown that the above definition implies that for any $x \in \mathbb{R}^n$ it holds the inequality

$$||Ax||_q \leq ||A||_{p \to q} ||x||_p.$$

An induced matrix norm is indeed a norm in the sense that it satisfies the three properties required above. We refer to the matrix norm $|| \cdot ||_{p,q}$s the (p, q)-norm. When $p = q$, we will simply refer to it as an p-norm and omit one of the subscripts in its notation.

An important operator norm is the spectral norm, where $p = q = 2$. In this case, it can be proved that the operator norm coincide with the maximum singular value of $A \in \mathbb{R}^{m \times n}$ (see below for more details on eigenvalues and singular values)

$$||A||_2 = ||A||_{2 \to 2} = \sqrt{\lambda_1(A^T A)} =: \sigma_1(A).$$

The default operator norm for this course will be the spectral norm, so when the subscript is omitted we will implicitly refer to this norm.

When $p = q = 1$, the operator norm reduces to

$$||A||_1 = ||A||_{1 \to 1} = \max_{j \in [n]} \sum_{i=1}^{m} |A_{ij}|.$$

This norm is also called maximum absolute column sum norm.

When $p = q = \infty$, the operator norm reduces to

$$||A||_\infty = ||A||_{\infty \to \infty} = \max_{i \in [m]} \sum_{j=1}^{n} |A_{ij}|.$$

This norm is also called the maximum absolute row sum norm.

An example of a matrix norm that is not induced by any norm is the Frobenius norm defined by

$$||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2} = ||\text{vec}(A)||,$$

in fact, you can think of this norm as the Euclidean norm of the vector obtained by flattening the matrix A into a vector.

2.5 Eigenvalues and Eigenvectors

Let $A \in \mathbb{R}^{n \times n}$. Then a nonzero vector $v \in \mathbb{R}^n$ is called an eigenvector of A if there exists a $\lambda \in \mathbb{C}$ (the complex field) for which

$$Av = \lambda v$$

The scalar λ is the eigenvalue corresponding to the eigenvector v. In general, real-valued matrices can have complex eigenvalues, but it is well known that all the eigenvalues of symmetric matrices are real. The eigenvalues of a symmetric $n \times n$ matrix A are denoted by

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A),$$

where $\lambda_1(A)$ is the maximum eigenvalue and $\lambda_n(A)$ is the minimum.

Definition 2.4. *A matrix U is said to be orthogonal when $U^T U = U U^T = Id$.*

One of the most useful results related to eigenvalues is the spectral decomposition theorem, which states that any symmetric matrix A has an orthonormal basis of eigenvectors.

Theorem 2.1 (Spectral Decomposition). *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix and let $v_1, \ldots, v_n$ be the eigenvectors corresponding to the eigenvalues $\lambda_1(A), \ldots, \lambda_n(A)$. Then $U = (v_1 | \ldots | v_n)$ is an orthonormal matrix and A can be decomposed as follows*

$$A = UDU^T, \quad \text{with } D := \mathrm{diag}(\lambda_1(A), \ldots, \lambda_n(A)).$$

A direct result of the spectral decomposition theorem is that the trace and the determinant of a matrix A can be expressed via its eigenvalues:

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i(A)$$

$$\det(A) = \prod_{i=1}^{n} \lambda_i(A)$$

Another important consequence of the spectral decomposition theorem is the bounding of the so-called Rayleigh quotient. For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, the Rayleigh quotient is defined as

$$R_A(x) := \frac{x^T A x}{||x||^2} \quad \forall x \neq 0,$$

and its upper and lower bounds are given by the following lemma.

Lemma 2.2. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then*

$$\lambda_n(A) \leq R_A(x) \leq \lambda_1(A) \quad \forall x \neq 0.$$

Moreover, we have the following corollary.

Corollary 2.1. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then*

$$\min_{x \neq 0} R_A(x) = \lambda_n(A),$$

and the eigenvector(s) of A corresponding to the minimal eigenvalue are the minimizers of this problem. In addition,

$$\max_{x \neq 0} R_A(x) = \lambda_1(A),$$

and the eigenvectors of A corresponding to the maximal eigenvalue are maximizers of this problem.

2.6 Singular values and Singular Vectors

Similar objects exist also for non-squared matrices $A \in \mathbb{R}^{m \times n}$.

Theorem 2.2 (Singular Value Decomposition). *Let $A \in \mathbb{R}^{m \times n}$. Then there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ and the uniquely defined real values*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_\ell \geq 0, \quad \text{with } \ell = \min\{m, n\},$$

such that

$$A = U \Sigma V^T \quad \text{with } \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_\ell) \in \mathbb{R}^{m \times n}.$$

The matrices $U = (u_1, \ldots, u_m)$ and $V = (v_1, \ldots, v_m)$ are composed respectively of the left and right singular vectors, thus we can decompose A also as follows

$$A = \sum_{j=1}^{\ell} \sigma_j u_j v_j^T.$$

If $\mathrm{rank}(A) = r$, then $\sigma_1, \geq \ldots, \geq \sigma_r > 0$ and $\sigma_{r+1} = \cdots = \sigma_\ell = 0$ so that the reduced singular value decomposition is

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^T.$$

Also, note that

$$A^T A = (U \Sigma V^T)^T U \Sigma V^T = V \Sigma^T U^T U \Sigma V^T = V \Sigma^T \Sigma V^T = V \mathrm{diag}(\sigma_1^2, \ldots, \sigma_\ell^2, 0, \ldots, 0) V^T$$

and

$$AA^T = U \Sigma V^T (U \Sigma V^T)^T = U \Sigma V^T V \Sigma^T U^T = U \Sigma \Sigma^T U^T = U \mathrm{diag}(\sigma_1^2, \ldots, \sigma_\ell^2, 0, \ldots, 0) U^T,$$

which implies that $\sigma_1^2, \ldots, \sigma_\ell^2$ are eigenvalues of $A^T A$ and AA^T, $u_1, \ldots, u_m$ are eigenvectors of AA^T and $v_1, \ldots, v_n$ are eigenvectors of $A^T A$. In other words the SVD can be obtained from the spectral decomposition of AA^T and $A^T A$. In particular,

$$\sigma_j = \sigma_j(A) = \sqrt{\lambda_j(A^T A)} = \sqrt{\lambda_j(AA^T)}$$

To conclude this section, the following function is called nuclear norm

$$||A||_* = \sum_{j=1}^\ell \sigma_j(A) \quad \text{with } \ell = \min\{m, n\},$$

and can actually proved to be a norm for $\mathbb{R}^{m \times n}$.

2.7 Basic Topological Concepts

Let us start with the definition of a ball.

Definition 2.5 (Open/Closed Ball). *The open ball with center $c \in \mathbb{R}^n$ and radius $r \in \mathbb{R}$ is denoted by $B(c, r)$ and defined by*

$$B(c, r) := \{x \in \mathbb{R}^n : ||x - c|| < r\}$$

The closed ball with center $c \in \mathbb{R}^n$ and radius $r \in \mathbb{R}$ is denoted by $B[c, r]$ and defined by

$$B[c, r] := \{x \in \mathbb{R}^n : ||x - c|| \leq r\}$$

The ball $B(c, r)$ for some arbitrary $r > 0$ is also referred to as a neighborhood of c. The first topological notion we define is that of an interior point of a set. These points have their neighborhood contained in the set.

Definition 2.6 (Interior Points). *Given a set $S \subseteq \mathbb{R}^n$, a point $c \in S$ is an interior point of S if there exists $r > 0$ for which $B(c, r) \subseteq S$.*

The set of all interior points of a given set S is called the interior of the set and is denoted by $\mathrm{int}(S)$:

$$\mathrm{int}(S) := \{x \in S : B(x, r) \subseteq S \text{ for some } r > 0\}.$$

Example 2.1. *Here are some examples*

$$\begin{aligned}
\mathrm{int}(\mathbb{R}_+^n) &= \mathbb{R}_{++}^n, \\
\mathrm{int}(B[c, r]) &= B(c, r) \quad \text{with } c \in \mathbb{R}^n, r \in \mathbb{R}_{++}, \\
\mathrm{int}([x, y]) &= (x, y) \quad \text{with } x, y \in \mathbb{R}^n, x \neq y.
\end{aligned}$$

Definition 2.7 (Open Set). *An open set is a set that contains only interior points. In other words, $S \subseteq \mathbb{R}^n$ is an open set if $\forall x \in S$ there exists $r > 0$ such that $B(x, r) \subseteq S$.*

Examples of open sets are $\mathbb{R}^n$, open balls, open line segments, and the positive orthant $\mathbb{R}_{++}^n$. A known result is that a union of any number of open sets is an open set and the intersection of a finite number of open sets is open.

Definition 2.8 (Closed Set). *A set $S \subseteq \mathbb{R}^n$ is said to be closed if it contains all the limits of convergent sequences of points in S, that is S is closed if for every sequence of points $\{x_k\} \subseteq S$ satisfying $\lim_{k \to \infty} x_k = x^*$, it holds that $x^* \in S$.*

Examples of closed sets are the closed balls, closed lines segments and the nonnegative orthant $\mathbb{R}_+$. The space $\mathbb{R}^n$ is both closed and open. An important and useful result states that level sets, as well as contour sets, of continuous functions are closed.

Proposition 2.1. *Let f be a continuous function over a closed set $S \in \mathbb{R}^n$. Then for any $\alpha \in \mathbb{R}$, the following sets are closed*

$$Lev(f, \alpha) := \{x \in S : f(x) \leq \alpha\}$$
$$Con(f, \alpha) := \{x \in S : f(x) = \alpha\}.$$

Definition 2.9 (Boundary Points). *Given a set $S \subseteq \mathbb{R}^n$, a boundary point of S is a point $x \in \mathbb{R}^n$ such that any neighborhood of x contains at least one point in S and at least one point in its complement S^c.*

The set of all boundary points of a set S is denoted by $\mathrm{bd}(S)$ and is called the boundary of S.

Example 2.2. *Some examples of boundary sets are*

$$\mathrm{bd}(B(c, r)) = \mathrm{bd}(B[c, r]) = \{x \in \mathbb{R}^n : ||x - c|| = r\} \quad \text{with } c \in \mathbb{R}^n, r \in \mathbb{R}_{++},$$
$$\mathrm{bd}(\mathbb{R}_{++}^n) = \mathrm{bd}(\mathbb{R}_+^n) = \{x \in \mathbb{R}_+^n : \exists x_i = 0\},$$
$$\mathrm{bd}(\mathbb{R}^n) = \emptyset.$$

Definition 2.10 (Closure of a Set). *The closure of a set $S \in \mathbb{R}^n$ is denoted by $\mathrm{cl}(S)$ and it is the union of S with its boundary*

$$\mathrm{cl}(S) = S \cup \mathrm{bd}(S).$$

Another equivalent definition of $\mathrm{cl}(U)$ is given by the smallest closet set containing S, i.e.,

$$\mathrm{cl}(S) := \bigcap \{T : S \subseteq T, T \text{ is closed}\}.$$

Example 2.3. *Some examples of closure are*

$$\mathrm{cl}(\mathbb{R}_{++}^n) = \mathbb{R}_+^n,$$
$$\mathrm{cl}(B(c, r)) = B[c, r] \quad \text{with } c \in \mathbb{R}^n, r \in \mathbb{R}_{++},$$
$$\mathrm{cl}((x, y)) = [x, y] \quad \text{with } x, y \in \mathbb{R}^n, x \neq y.$$

Definition 2.11 (Boundedness). *A set S is bounded when there exists a $M > 0$ such that $S \subseteq B(0, M)$.*

Definition 2.12 (Compactness). *A set S is compact when it is bounded and closed.*

Examples of compact sets are closed balls and line segments. The positive orthant is not compact since it is unbounded, and open balls are not compact since they are not closed.

2.8 Differentiability

Definition 2.13 (Directional Derivative). *Given a function f defined on a set $S \subseteq \mathbb{R}^n$, a point $x \in \mathrm{int}(S)$ and a direction $d \in \mathbb{R}^n$, with $d \neq 0$, the directional derivative of f at x along d is the below limit*

$$\lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t},$$

when it exists. In this case, the directional derivative is denoted by $f'(x, d)$.

Definition 2.14 (Partial Derivative). *For any $i \in [n]$, the directional derivative of f at x along the direction e_i is called the i-th partial derivative and it is denoted by*

$$\frac{\partial f}{\partial x_i}(x) := \lim_{t \to 0^+} \frac{f(x + te_i) - f(x)}{t}.$$

Definition 2.15 (Gradient). *If all partial derivatives of a function f exist at a point $x \in \mathbb{R}^n$, then the gradient of f at x is defined by the column vector consisting of all partial derivatives of f:*

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

A function f defined on an open set $S \subseteq \mathbb{R}^n$ is called continuously differentiable (or smooth) over S, and we write $f \in \mathrm{C}^1(S)$, if all the partial derivatives exist and are continuous on S.

An important formula to compute the directional derivative is the following

$$f'(x,d) = \nabla f(x)^T d \quad \forall x \in S, d \in \mathbb{R}^n.$$

The partial derivatives $\frac{\partial f}{\partial x_i}$ are themselves real-valued functions that can be partially differentiated. The second order (i,j)-th partial derivative of f at $x \in S$ (if exists) is defined by

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

A function f defined on an open set $S \subseteq \mathbb{R}^n$ is called twice continuously differentiable over S, and we write $f \in \mathrm{C}^2(S)$, if all the second order partial derivatives exist and are continuous over S. When f is twice continuously differentiable, the second order partial derivatives are symmetric, meaning that for any $i \neq j$ and any $x \in S$

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

Definition 2.16 (Hessian). *Assuming $f \in \mathrm{C}^2(S)$, we can compute the Hessian of f at a point $x \in S$ as the symmetric $n \times n$ matrix*

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2}{\partial x_1^2}(x) & \cdots & \frac{\partial^2}{\partial x_1 x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n x_1}(x) & \cdots & \frac{\partial^2}{\partial x_n^2}(x) \end{pmatrix}.$$

In the following we list a few important theorems that are consequence of Taylor's approximation theorem and that might be useful in the rest of the course.

Theorem 2.3 (Linear Approximation). *Let $f \in \mathrm{C}^1(S)$, with $S \subseteq \mathbb{R}^n$ open, and let $x \in S, r > 0$ satisfy $B(x,r) \subseteq S$. Then for any $y \in B(x,r)$, we have*

$$f(y) = f(x) + \nabla f(x)^T (y-x) + o(||y-x||),$$

where $o(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is a one dimensional function satisfying $\frac{o(t)}{t} \to 0$ as $t \to 0^+$.

Theorem 2.4 (Quadratic Approximation). *Let $f \in \mathrm{C}^2(S)$, with $S \subseteq \mathbb{R}^n$ open, and let $x \in S, r > 0$ satisfy $B(x,r) \subseteq S$. Then for any $y \in B(x,r)$, we have*

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

Theorem 2.5 (Mean Value Theorem). *Let $f \in \mathrm{C}^1(S)$, with $S \subseteq \mathbb{R}^n$ open, and let $x \in S, r > 0$ satisfy $B(x,r) \subseteq S$. Then for any $y \in B(x,r)$, there exists a $c \in [x,y]$ such that*

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

Theorem 2.6. *Let $f \in \mathrm{C}^2(S)$, with $S \subseteq \mathbb{R}^n$ open, and let $x \in S, r > 0$ satisfy $B(x,r) \subseteq S$. Then for any $y \in B(x,r)$, there exists a $c \in [x,y]$ such that*

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

3 Optimality Condition for Unconstrained Optimization

3.1 Global and Local Optima

Despite our main interest in this section is to discuss minimum and maximum points of a function over the entire space (unconstrained), we here still present the more general definition of global minimum and maximum points of a function over a given set.

Definition 3.1 (Global Minimum and Maximum). *Let $f : S \to \mathbb{R}$, then*

1. $x^* \in S$ is a global minimum point of f over S if $f(x^*) \leq f(x) \; \forall x \in S$;

2. $x^* \in S$ is a *strict global minimum point* of f over S if $f(x^*) < f(x) \ \forall x \in S, \ x \neq x^*$;

3. $x^* \in S$ is a *global maximum point* of f over S if $f(x^*) \geq f(x) \ \forall x \in S$;

4. $x^* \in S$ is a *strict global maximum point* of f over S if $f(x^*) > f(x) \ \forall x \in S, \ x \neq x^*$.

The set S on which the optimization of f is performed is also called the **feasible set**, and any point $x \in S$ is called **feasible solution**. It is common to refer to a global minimum point as a **minimizer** or a global minimizer.

Despite they might not be attained, the **minimal** and **maximal** value of f over S is the classical notation to refer to the infimum and supremum of f over S. These values are always **unique**, while there might exist many global minima/maxima attaining the same value of f, i.e. $f(x^*)$. The sets of all global minimizers/maximizers of f over S are denoted by

$$\operatorname{argmin}\{f(x) : x \in S\} \qquad \operatorname{argmax}\{f(x) : x \in S\}.$$

A vector $x^* \in S$ is called a global **optimum** of f over S if it is either a global minimum or a global maximum.

Our main task will usually be to find and study global minimum or maximum points; however, most of the theoretical results only characterize local minima and maxima which are optimal points with respect to a neighborhood of the point of interest.

Definition 3.2 (Local Minima and Maxima). *Let $f : S \to \mathbb{R}$, then*

1. $x^* \in S$ is a *local minimum point* of f over S if there exists $r > 0$ for which
   $f(x^*) \leq f(x) \ \forall x \in S \cup B(x^*, r)$;

2. $x^* \in S$ is a *strict local minimum point* of f over S if there exists $r > 0$ for which
   $f(x^*) < f(x) \ \forall x \in S \cup B(x^*, r), \ x \neq x^*$;

3. $x^* \in S$ is a *local maximum point* of f over S if there exists $r > 0$ for which
   $f(x^*) \geq f(x) \ \forall x \in S \cup B(x^*, r)$;

4. $x^* \in S$ is a *strict local maximum point* of f over S if there exists $r > 0$ for which
   $f(x^*) > f(x) \ \forall x \in S \cup B(x^*, r), \ x \neq x^*$.

Notice that a global minimum (maximum) point is also a local minimum (maximum) point.

A well-known result (**Fermat's theorem**) is that for a one-dimensional function f defined and differentiable over an interval (a, b), if a point $x^* \in (a, b)$ is a local optimum, then $f'(x^*) = 0$. The multidimensional extension of this result states that the gradient is zero at local optima. We refer to such an optimality condition as a first order optimality condition, as it is expressed in terms of the first order derivatives.

3.2 First Order Optimality Conditions

Theorem 3.1 (First Order Optimality Condition). *Let $f : S \to \mathbb{R}$ be a function defined on $S \subseteq \mathbb{R}^n$. Suppose that $x^* \in \operatorname{int}(S)$ is a local optimum and that all partial derivatives of f exist at x^*. Then $\nabla f(x) = 0$.*

Proof. Let $i \in [n]$ and consider the one-dimensional function $g(t) = f(x^* + te_i)$. Note that g is differentiable at $t = 0$ and by definition $g'(0) = \frac{\partial f}{\partial x_i}(x^*)$. Since x^* is a local optimum of f, if follows that $t = 0$ is a local optimum of g, which, by Fermat's Theorem, directly implies that $g'(0) = 0$ and thus also $\frac{\partial f}{\partial x_i}(x^*) = 0$. Now, since the argument can be repeated for all $i \in [n]$, the result $\nabla f(x^*) = 0$ follows. $\quad \square$

Theorem 3.1 presents a necessary optimality condition: the gradient vanishes at all local optimum points, which are interior points of the domain of the function. However, the **reverse claim is not true** as there could be points which are not local optimum points, whose gradient is zero. The points in which the gradient vanishes are called **stationary points** and they are the only candidates for local optima among the points in the interior of the domain of the function.

Definition 3.3 (Stationary Points). *Let $f : S \to \mathbb{R}$ be a function defined on a set $S \subseteq \mathbb{R}^n$. Suppose that $x^* \in \operatorname{int}(S)$ and that f is differentiable in a neighborhood of x^*. Then x^* is called stationary point of f if $\nabla f(x^*) = 0$.*

3.3 Classification of Matrices

Definition 3.4 (Positive Semidefiniteness)**.** *A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive semidefinite, denoted by $A \succcurlyeq 0$, if $x^T A x \geq 0$ for every $x \in \mathbb{R}^n$.*

Definition 3.5 (Positive Definiteness)**.** *A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive definite, denoted by $A \succ 0$, if $x^T A x > 0$ for every $x \in \mathbb{R}^n, x \neq 0$.*

In this course, a positive definite or semidefinite matrix is always assumed to be symmetric. The same definitions can be given for negative (semi)definite matrices. Notice that a matrix A is negative (semi)definite if and only if $-A$ is positive (semi)definite. A more complicated matrix is instead the following.

Definition 3.6 (Indefiniteness)**.** *A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is indefinite, if there exist $x, y \in \mathbb{R}^n$ such that $x^T A x > 0$ and $y^T A y < 0$.*

Indefinite matrices are easier to recognize than positive (semi)definite ones, as matrices that contain positive and negative values on the diagonal are indefinite. To better characterize the latter, we need to look the eigenvalues of the matrix.

Theorem 3.2 (Eigenvalue characterization theorem)**.** *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Then*

1. *A is positive definite if and only if all its eigenvalues are positive,*

2. *A is positive semidefinite if and only if all its eigenvalues are nonnegative,*

3. *A is indefinite if and only if it has at least one positive eigenvalue and at least one negative eigenvalue.*

3.4 Second Order Optimality Conditions

Theorem 3.3 (Necessary Second Order Optimality Condition)**.** *Let $f : S \to \mathbb{R}$ be a function defined on a open set $S \subseteq \mathbb{R}^n$. Suppose that $f \in C^2(S)$ and that x^* is a stationary point. Then the following hold:*

(a) *If x^* is a local minimizer of f over S, then $\nabla^2 f(x^*) \succcurlyeq 0$;*

(b) *If x^* is a local maximizer of f over S, then $\nabla^2 f(x^*) \preccurlyeq 0$*

Proof. (a) Since x^* is a local minimizer, there exists a ball $B(x^*, r) \subseteq U$ for which $f(x) \geq f(x^*) \ \forall x \in B(x^*, r)$. Let $d \in \mathbb{R}^n$ be a nonzero vector. For any $0 < \alpha < \frac{r}{||d||}$ we have $x_\alpha^* := x^* + \alpha d \in B(x^*, r)$, which means that for any such α

$$f(x^*) \leq f(x_\alpha^*). \tag{2}$$

On the other hand from Theorem 2.6 it follows that there exists a vector $z_\alpha \in [x^*, x_\alpha^*]$ such that

$$f(x_\alpha^*) - f(x^*) = \nabla f(x^*)^T (x_\alpha^* - x^*) + \frac{1}{2} (x_\alpha^* - x^*)^T \nabla^2 f(z_\alpha)(x_\alpha^* - x^*).$$

Since x^* is a stationary point of f, and by the definition of x_α^*, the latter equation reduces to

$$f(x_\alpha^*) - f(x^*) = \frac{\alpha^2}{2} d^T \nabla^2 f(z_\alpha) d,$$

which combined with (2) yields that for any $\alpha \in (0, \frac{r}{||d||})$ the inequality $d^T \nabla^2 f(z_\alpha) d \geq 0$ holds. Finally, using the fact that $z_\alpha \to x^*$ as $\alpha \to 0^+$, and the continuity of the Hessian, we obtain that $d^T \nabla^2 f(x^*) d \geq 0$. Since the latter inequality holds for any $d \in \mathbb{R}^n$, the desired result is established.

The proof of (b) follows directly from employing the result of part (a) on the function $-f$. □

Theorem 3.4 (Sufficient Second Order Optimality Condition)**.** *Let $f : S \to \mathbb{R}$ be a function defined on a open set $S \subseteq \mathbb{R}^n$. Suppose that $f \in C^2(S)$ and that x^* is a stationary point. Then the following hold:*

(a) *If $\nabla^2 f(x^*) \succ 0$, then x^* is a strict local minimizer of f over S;*

(b) *If $\nabla^2 f(x^*) \prec 0$, then x^* is a strict local maximizer of f over S.*

Definition 3.7 (Saddle Points)**.** *Let $f : S \to \mathbb{R}$ be a function defined on a open set $S \subseteq \mathbb{R}^n$. Suppose that $f \in C(S)$. A stationary point x^* is a saddle point of f over S when it is neither a local minimizer nor a local maximizer of f over S.*

Theorem 3.5 (Sufficient Condition for a Saddle Point)**.** *Let* $f : S \to \mathbb{R}$ *be a function defined on a open set* $S \subseteq \mathbb{R}^n$. *Suppose that* $f \in C^2(S)$ *and that* x^* *is a stationary point. If* $\nabla^2 f(x^*)$ *is an indefinite matrix, then* x^* *is a saddle point for* f *over* S.

Theorem 3.6 (Sufficient Condition for a Global Minimizer)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *with* $f \in C^2(\mathbb{R}^n)$ *and* $\nabla^2 f(x) \succcurlyeq 0$ *for any* $x \in \mathbb{R}^n$. *Let* x^* *be a stationary point of* f. *Then* x^* *is a global minimizer of* f.

Proof. By the Mean Value Theorem (Theorem 2.5), it follows that for any $x \in \mathbb{R}^n$, there exists a vector $z_x \in [x^*, x]$ for which

$$f(x) - f(x^*) = \frac{1}{2}(x - x^*)^T \nabla^2 f(z_x)(x - x^*).$$

Since $\nabla^2 f(x) \succcurlyeq 0$, we have that $f(x) \geq f(x^*)$, concluding the proof. $\qquad \square$

Remark 3.1. *At a first sight, this result seems asymmetric, what about global maximizers? Maximization problems can be easily transformed in minimization problems*

$$\min_x f(x) = -\max_x -f(x).$$

3.5   Sufficient Conditions for Existence of Optima

Another important issue is the one of deciding on whether a function actually has a global minimizer or maximizer. This is the issue of **existence**. A very well known result is due to Weierstrass, stating that a continuous function attains its minimum and maximum over a compact set.

Theorem 3.7 (Weierstrass)**.** *Let* f *be a continuous function defined over a nonempty and compact set* $C \subseteq \mathbb{R}^n$. *Then there exists a global minimum point of* f *over* C *and a global maximum point of* f *over* C.

When the underlying set is not compact, the Weierstrass theorem does not guarantee the attainment of the solution, but certain properties of the function f can imply the existence of the solution even in the noncompact setting. One example of such a property is coerciveness.

Definition 3.8 (Coerciveness)**.** *Let* f *be a continuous function defined over* $\mathbb{R}^n$. *The function is said to be coercive when*

$$\lim_{||x|| \to \infty} f(x) = \infty.$$

Theorem 3.8 (Existence under Coerciveness)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a continuous and coercive function and let* $S \subseteq \mathbb{R}^n$ *be a nonempty closed set. Then* f *has a global minimum point over* S.

Proof. Let $x_0 \in S$ be an arbitrary point. Since the function is coercive, it follows that there exists a $M > 0$ such that

$$f(x) > f(x_0) \quad \text{for any } x \text{ such that } ||x|| > M. \tag{3}$$

Since any global minimizer x^* of f over S satisfies $f(x^*) \leq f(x_0)$, it follows from (3) that the set of global minimizers of f over S is the same as the set of global minimizers of f over $S \cap B[0, M]$. The set $S \cap B[0, M]$ is compact and nonempty and thus by the Weierstrass theorem, there exists a global minimizer of f over $S \cap B[0, M]$ and hence also over S. $\qquad \square$

References

[1] Harry M Markowits. Portfolio selection. *Journal of finance*, 7(1):71–91, 1952.