Chapter 2

Convex Analysis

Convexity is a central concept in nonlinear programming.

2.1 Convex set

Definition 2.2 *A* convex combination *of $p \geq 1$ vectors $x^1, \ldots, x^p$ is a vector of the form*

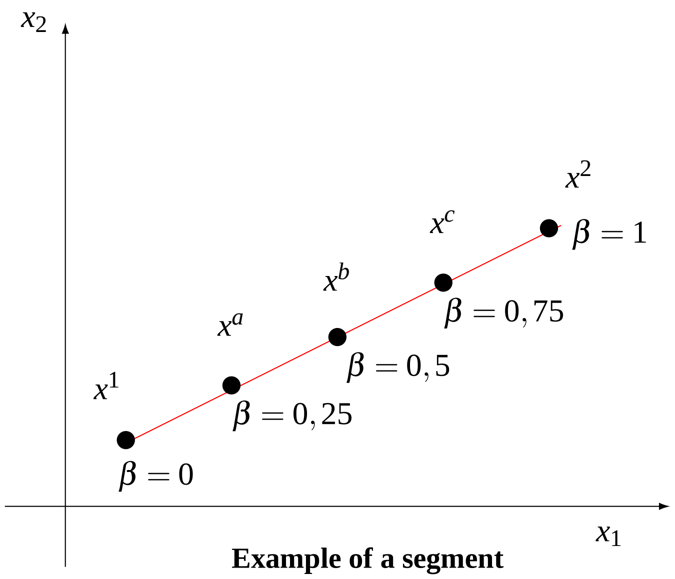
$$y = \sum_{i=1}^{p} \lambda_i x^i \qquad \sum_{i=1}^{p} \lambda_i = 1 \qquad \lambda_i \geq 0 \;\; i = 1, \ldots, p.$$

Definition 2.3 *Let x e y in $\Re^n$. The convex combination of the two points*

$$z = (1 - \beta)x + \beta y,$$

being β in the interval $[0, 1]$ is called the closed segment *with extremes x e y and it is denoted as $[x, y]$.*

Example 2.4 In figure below is plot the segment with extremes $x^1 = (1,1)^T$ and $x^2 = (8,5)^T$. For $\beta = 0$ the corresponding point is x^1, while with $\beta = 1$ we get x^2; point denoted in the picture as x_a, x_b e x_c correspond to β equal respectively to 0.25, 0.5 e 0.75.



Example of a segment

Definition 2.5 *A set $X \subseteq \Re^n$ is* convex *if*

$$(1-\beta)x + \beta y \in X \qquad \forall \quad x,y \in C \text{ and } \beta \in [0,1]$$

Alternatively we can say that a set X is convex if for any $x,y \in C$ we have $[x,y] \subseteq X$.

Theorem 2.6 *The intersection of two convex sets is convex.*

Proof. Let $X_1, X_2 \subseteq \Re^n$ convex sets and $X = X_1 \cap X_2$. Let x and y belonging to X; by definition both $x,y \in X_1$ and $x,y \in X_2$. Since X_1 and X_2 are convex then $[x,y] \subseteq X_1$ and $[x,y] \subseteq X_2$ so that $[x,y] \subseteq X$ and the proof is completed.

Theorem 2.7 *The intersection of any finite number of convex sets is convex.*

Let's consider the linear equality

$$a_1 x_1 + \quad \dots \quad + a_n x_n = b$$

or in matricial notation $a^T x = b$, with $a \in \Re^n$ e $b \in \Re$.

Definition 2.8

$$H = \{x \in \Re^n : a^T x = b\}$$

is called hyperplane defined by the equation $a^T x = b$. The sets

$$S^{\leq} = \{x \in \Re^n : a^T x \leq b\}$$
$$S^{\geq} = \{x \in \Re^n : a^T x \geq b\}$$

are closed semispace defined by $a^T x \leq b$ e $a^T x \geq b$.

In $\Re^2$ an hyperplane is a line; in $\Re^3$ it is a plane. Semispace represent set of points that ”lie” on the same side of the hyperplane.

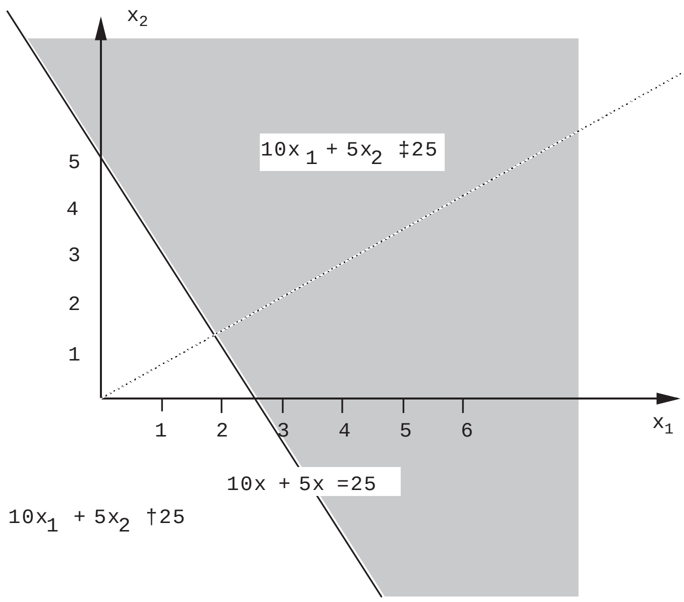


Figure 2.1: Line and semispaces.

We have that

$$H \subset S^{\geq}, \qquad H \subset S^{\leq}, \qquad S^{\geq} \cap S^{\leq} = H.$$

Semispaces and hyperplanes are convex sets.

Theorem 2.9 *A closed semispace is a convex set.*

Proof: without loss of generality we prove for $S^{\leq} = \{x \in \Re^n : a^T x \leq b\}$. Let $x, y \in S^{\leq}$. Let $z = \beta x + (1-\beta)y$ with $0 \leq \beta \leq 1$. Since both x and y belong $S^{\leq}$ they satisfy $a^T x \leq b$, $a^T y \leq b$. Since both coeffciients $\beta \geq 0$ and $1 - \beta$ $ge0$ we get

$$a^T(\beta x + (1-\beta)y) = \beta a^T x + (1-\beta)a^T y \leq \beta b + (1-\beta)b = b$$

so that $a^T z \leq b$ which proves $z \in S^{\leq}$.

Using theorems 2.9 and 2.6 we easily get convesso.

Theorem 2.10 *An hyperplane is a convex set.*

2.11 Convex and concave functions

Definition 2.12 *A function $f(x)$ is* convex *on a convex set $\mathscr{C}$ satisfies*

$$f((1-\beta)y+\beta z) \leq (1-\beta)f(y)+\beta f(z), \quad \beta \in [0,1] \text{ for all } y,z \in \mathscr{C}. \tag{2.1}$$

A function $f(x)$ is strictly convex *if for all $y,z \in \mathscr{C}$, with $y \neq z$, it holds*

$$f((1-\beta)y+\beta z) < (1-\beta)f(y)+\beta f(z), \quad \beta \in (0,1).$$

Definition 2.13 *A function $f(x)$ is* concave *on a convex set $\mathscr{C}$ satisfies*

$$f((1-\beta)y+\beta z) \geq (1-\beta)f(y)+\beta f(z), \quad \beta \in [0,1] \text{ for all } y,z \in \mathscr{C}. \tag{2.2}$$

A function $f(x)$ is strictly concave *if for all $y,z \in \mathscr{C}$, with $y \neq z$, it holds*

$$f((1-\beta)y+\beta z) > (1-\beta)f(y)+\beta f(z), \quad \beta \in (0,1).$$

A function $f(x)$ is *(strictly) concave* on a convex set $\mathscr{C}$ if the function $-f(x)$ is *(strictly) convex* on $\mathscr{C}$.

2.14 Optimization problems

We have the following classification

- Convex optimization problems: a minimum problem with convex objective function and convex feasible set (or maximum problem with concave objective function and convex feasible set)
- Concave optimization problems: a minimum problem with concave objective function and convex feasible set (or maximum problem with convex objective function and convex feasible set)
- general problems, when neither of the two conditions above is satisfied.

An optimization (min or max) problem with linear objective function over a convex set S is both a convex and a concave optimization problem.

Hence we have

A linear programming problems is both a convex and a concave optimization problem.

Proposition 2.15 **[Absence of local solution of a convex problem]** *Let $S \subseteq R^n$ a convex set and f a convex function on S. Then, either the problem*

$$\begin{array}{ll} \min & f(x) \\ & x \in S \end{array}$$

has no solution, or a local minimum of f is also a global minimum. If in addition f is strictly convex, then there exists at most one global minimum of f.

Proof. Assume that a global solution x^* exists. By contraddiction we assume that there exists a local solution $\hat{x}$ which is not global. Hence it must hold that $f(x^*) < f(\hat{x})$.

By defintion of local solution we have

$$f(\hat{x}) \leq f(x) \quad \forall\, x \in S \cap \mathscr{N}(\hat{x};\rho).$$

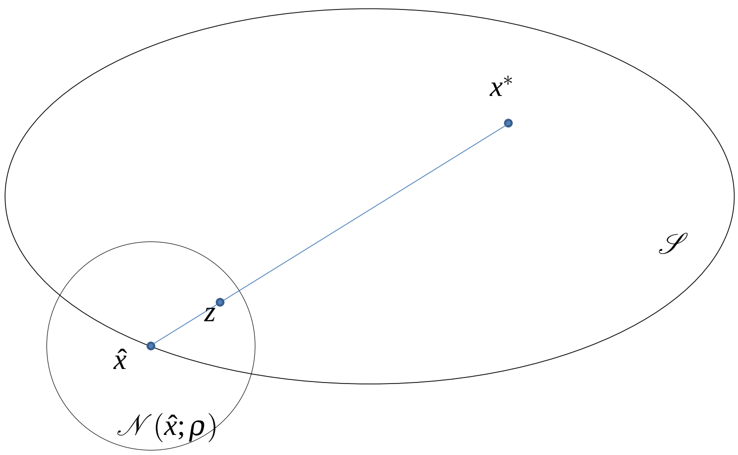


Figure 2.2: Geometric representation of the proof of Theorem 2.15 the position of $\hat{x}, x^*$ can be changed without affecting the proof

Let us consider the segment $[\hat{x}, x^*] \in S$. By convexity of f, we have:

$$f((1-\beta)\hat{x} + \beta x^*) \leq (1-\beta) f(\hat{x}) + \beta f(x^*) = f(\hat{x}) + \beta(f(x^*) - f(\hat{x})), \text{for all } \beta \in [0,1].$$

The last term $\beta(f(x^*) - f(\hat{x}))$ is < 0, for all $\beta \in (0,1]$, and it is zero only for $\beta = 0$. Hence for all the points in $z \in (\hat{x}, x^*]$ we get $f(z) < f(\hat{x})$, that means there exists no neighborhood of radius $\rho > 0$ in which $\hat{x}$ satisfies the definition of local solution.

Theorem 2.16 **[No internal solution for concave problems]** *Let $S \subseteq R^n$ be a closed convex set and f a convex function on S. Assume that f is not constant on S. Then, either the problem*

$$\begin{array}{ll} \min & f(x) \\ & x \in S \end{array}$$

has no solution, or every global solution is on the boundary of the set S.

Proof. Let x^* be any optimal solution. Since by assumption f is not constant on S, there must exist a point $\hat{x} \in S$ such that

$$f(\hat{x}) > f(x^*) = \min_{x \in S} f(x).$$

Consider an interior point $z \in S$. By defintion there exists a spherical neighborhood $B(z;\rho)$ centered in z with radius $\rho > 0$ fully contained in S. On the line joining $\hat{x}$ and z we fixed a point $y \in B(z;\rho) \subseteq S$ with $y \neq z$ such that $z \in [\hat{x}, y]$. Hence there exists a λ con $\lambda \in [0,1)$ such that

$$z = (1-\lambda)\hat{x} + \lambda y.$$

By concavity of f and using $f(\hat{x}) > f(x^*)$, $f(y) \geq f(x^*)$ and $1-\lambda > 0$ (because $y \neq z$), we get:

$$f(z) \geq (1-\lambda) f(\hat{x}) + \lambda f(y) > (1-\lambda) f(x^*) + \lambda f(x^*) = f(x^*).$$

Hence for any z we get $f(z) > f(x^*)$ so that no interior point can be a solution. □

As regard Linear programming problems we get as a consequence the following result

Theorem 2.17 *If a linear programming (LP) problem admits a solution, then the solutions are all on the boundary of the feasible polyhedron.*

2.18 Characterization of continuously diffentiable convex functions

Theorem 2.19 (Necessary and sufficient condition for convexity) *Let C be an open convex set and $f : C \to R$ with continuous ∇f over C. The function f is convex on C if and only if for all $x, y \in C$ it holds:*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x). \tag{2.3}$$

Further f is strictly convex on C if and only if for all $x, y \in C$ con $y \neq x$ it holds

$$f(y) > f(x) + \nabla f(x)^T (y - x). \tag{2.4}$$

Theorem 2.20 (Necessary and sufficient condition for convexity) *Let C be an open convex set and $f : C \to R$ with continuous $\nabla^2 f$ over C. The function f is convex on C if and only if for all $x \in C$ $\nabla^2 f(x)$ is positive semidefinite.*

Theorem 2.21 (Sufficient condition for strict convexity) *Let C be an open convex set and $f : C \to R$ with continuous $\nabla^2 f$ over C. If $\nabla^2 f(x)$ is positive definite. then the function f is strictly convex on C.*

2.21.1 Quadratic form and quadratic functions

A *quadratic function* is

$$q(x) = \frac{1}{2}x^T Qx + c^T x,$$

where Q is a square symmetric matrix $n \times n$ with elements $q_{ij} = q_{ji},\ i, j = 1, \dots, n$. we have

$$\nabla f(x) = Qx \qquad\qquad \nabla^2 f(x) = Q.$$

Given a square symmetric matrix Q the associated *quadratic form* is

$$x^T Qx = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_i x_j. \tag{2.5}$$

The quadratic form $x^T Qx$ is said to be

- *positive definite* if $x^T Qx > 0\ \forall x \in \Re^n, x \neq 0$;

- *positive semidefinite* if $x^T Qx \geq 0\ \forall x \in \Re^n$;

- *indefinite* if there are x such that $x^T Qx > 0$ and $x^T Qx < 0$.

Corrisponding the matrix Q associated to the quadratic form is said to be positive definite, positive semidefinite and indefinite respectively.

The quadratic form $x^T Qx$ is said to be negative (semi)definite if $-x^T Qx$ is positive (semi)definite.

We have the following theorems.

Proposition 2.22 *A quadratic function $q(x)$ is convex on $\Re^n$ if and only if*

$$\frac{1}{2}y^T Qy \geq 0, \quad \text{per forall } y \in \Re^n;$$

the function $q(x)$ is strictly convex on $\Re^n$ if and only if

$$\frac{1}{2}y^T Qy > 0, \quad \text{per forall } y \in \Re^n.$$

Criteria for checking positive definiteness of a matrix. To check if a symmetric matrix A is positive definite the following criterion can be used.

Criterion for positive definiteness Let $A_k, k = 1\ldots,n$ the n north-west principal minors A, i.e. the n sub matrices with elements $a_{ij}, i, j = 1,\ldots,k$, obtained by eliminating from A the last $n-k$ rows and columns.

Let $\det A_k$ be the determinant of A_k.

- A is positive definite if and only if $\det A_k > 0$, for all $k = 1,\ldots,n$.

We remark that if A is positive semidefinite then $\det A_k \geq 0$, but vice versa is not true. Indeed consider the matrix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 2a_{22} \end{pmatrix}$$

with $a_{22} < 0$ and the quadratic

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

with $a_{11} = a_{12} = 0$. The principal minors A_1 e A_2 have determinant equal to zero so that they satisfies $\det A_k \geq 0$, but A is negative semidefinite!!!

Checking whether a matrix $A = \{a_{ij}\}_{i,j=1,\ldots,n}$ is positive semidefinite is much more computationally onerous. We have the following criterion.

Criterion for positive semidefiniteness let $J^k = \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\}$ and the $|J^k| \times |J^k|$ square matrix $D_{J^kJ^k}$ obtained from A eliminating the h-th column and row for all $h \notin J^k$. These matrix $D_{J^kJ^k} = \{a_{ij}\}_{i=i_1,\ldots,i_k,\ j=i_1,\ldots,i_k}$ are the principal minors of A. Let $\det D_{i_k,j_k}$ idenotes the determinant of $D_{J^kJ^k}$.

- A is positive semidefinite if and only if $D_{J^kJ^k} \geq 0$ for all $J^k \subset \{1,\ldots,n\}$.

Example 2.23 The strictly convex quadratic function

$$4x_1^2 + 8x_2^2 - x_1$$

is plot in the figure 2.3.

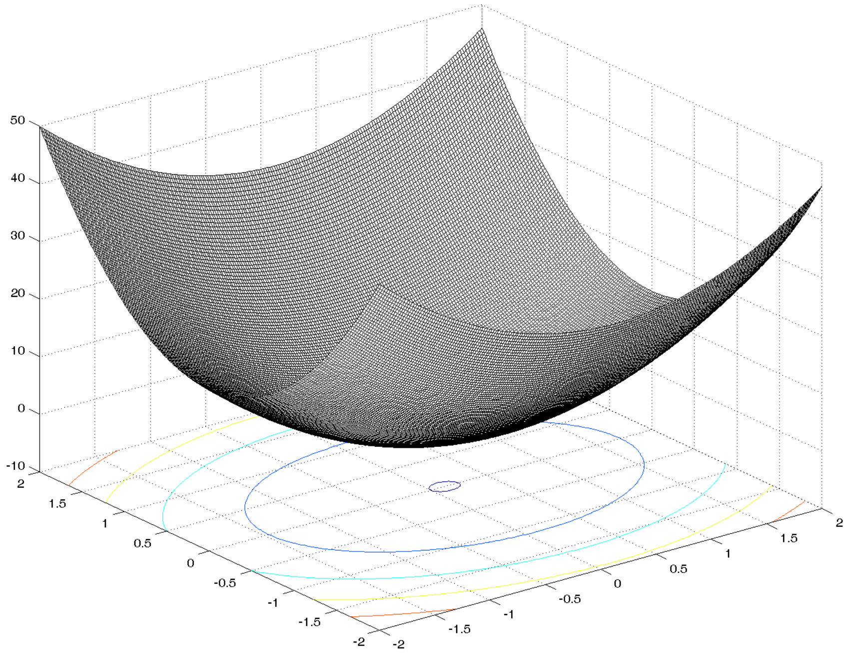


Figure 2.3: Plot of the convex quadratic function Example 2.23.