Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

INFORMATION ENGINEERING DEPARTMENT

COMPUTER ENGINEERING - AI & ROBOTICS

OPERATIONS RESEARCH 1 2022/2023

Francesco Crisci 2076739

Convex Programming

Definition 1.1.1

Given $x, i \in R^n$, the point $z := \lambda x + (1-\lambda)y$ is said to be a convex combination of x,y for all $\lambda \in [0,1]$. The combination is said to be strict if $0 < \lambda < 1$.
More generally, the convex combination of k points $x^1, ..., x^k \in R^n$ is defined as $\Sigma_{i=1}^k \lambda_i x^i$, with $\lambda_i, ..., \lambda_k \geq 0$ and $\Sigma_{i=1}^k \lambda_i = 1$.

Definition 1.1.2

A set $X \subseteq R^n$ is said to be convex if $\forall x, y \in X$ we have that X contains all the convex combinations of x and y, i.e : $z := [\lambda x + (1-\lambda)y], \forall \lambda \in [0,1]$.

Proposition 1.1.1

The intersection of two convex set $A, B \in R^n$ is still a convex set.
Proof:
Given $x, y \in A \cap B$, for all $\lambda \in [0,1]$ we have $z := \lambda x + (1-\lambda)y \in A$ by convexity of A, $z \in B$ for convexity of B, hence $z \in A \cap B$, as requested.

Definition 1.1.3

A function $f : X \to R$ defined on a convex set $X \subseteq R^n$ is said to be convex if $\forall x, y \in X$ and $\forall \lambda \in [0,1]$ we have that $f(z) \leq \lambda f(x) + (1-\lambda) f(y)$, where $z = \lambda x + (1-\lambda)y$.

Theorem 1.1.1

Let $X = \{x \in R^n : g_i(x) \leq 0, i = 1, ..., m\}$. If for all $i \in \{1, ..., m\}$ the functions $g_i : R^n \to R$ are convex, then the set X is convex.
Proof:
Clearly $X = \cap_{i=1}^m X_i$, where $X_i := \{x \in R^n : g_i(x) \leq 0\}$. By proposition 1.1.1, it is then sufficient to prove that each set X_i is convex. Indeed, given any two elements x and y of X_i and a generic point $x = \lambda x + (1-\lambda)y, \lambda \in [0,1]$, by the convexity hypothesis of the function g_i we can write $g_i(z) = g_i(\lambda x + (1-\lambda)y) \leq \lambda g_i(x) + (1-\lambda) g_i(y) \leq 0$, where the latter inequality is valid since $g_i(x) \leq 0, g_i(y) \leq 0$, and $0 \leq \lambda \leq 1$. It follows that $g_i(z) \leq 0$, hence $z \in X_i$. Given the arbitrariness of x,y and z, one thus has that X_i is convex, as requested.

Theorem 1.1.2

Consider a convex programming problem, i.., a problem $\min\{f(x) : x \in X\}$ where $X \in R^n$ is a convex set and $f : X \to R$ is a convex function. Every locally optimal solution is also a globally optimal solution.
Proof:
Let $\tilde{x}$ be any locally optimal solution. By the lcal optimum definition, there exists then $\epsilon > 0$ such that $f(\tilde{x}) \leq f(z)$

for all $z \in I_\epsilon(\tilde{x}) := \{x \in X : |x - \tilde{x}|| \leq \epsilon\}$. We have to prove that $f(\tilde{x}) \leq f(y)$ for all $y \in X$. Given any $y \in X$, consider the point z belonging to the segment that connects $\tilde{x}$ to y and define as $z := \lambda\tilde{x} + (1-\lambda)y$, where $\lambda < 1$ is chose very close to the value 1 so that $z \in I_\epsilon(\tilde{x})$ and hence $f(\tilde{x}) \leq f(z)$. By the convexity hypothesis of f it follows that $f(\tilde{x}) \leq f(z) = f(\lambda\tilde{x} + (1-\lambda)y) \leq \lambda f(\tilde{x}) + (1-\lambda)f(y)$, from which, dividing by $(1-\lambda) > 0$, we obtain $f(\tilde{x}) \leq f(y)$, as requested.

Simplex Algorithm

Definition 4.1.1

The sets $\{x \in R^n : \alpha^T x \leq \alpha_0\}$ and $\{x \in R^n : \alpha^T = \alpha_0\}$ are called affine half.space and hyperplane, respectively, induced by (α, α_0).

Definition 4.1.2

A (convex) polyhedron is defined as the intersection of a finite number of affin ehalf-spaces and hyperplanes. The sets of feasible solutions of Linear Programming problems are hence polyhedra.

Definition 4.1.3

A bounded (i.e. there exists $M > 0$ such that $||x|| \leq M$ for all $x \in P$) polyhedron P is called polytope.

Definition 4.1.4

A point x of a polyhedron P is said to be an extreme point or a vertex of P if it cannot be expressed as a strict convex combination of other two points of the polyhedron, i.e, if there exist no $y, z \in P, y \neq z$ and $\lambda \in (0, 1)$ such that $x = \lambda y + (1 - \lambda)z$.

Theorem 4.1.1 Minkowski-Weyl theorem

Every point of a polytope can be obtained as the convex combination of its vertices.

Theorem 4.1.2

If the set P of the feasible solutions of the linear programming problem $min\{c^T x : x \in P\}$ is bounded, then there exists at least one optimal vertex of P.
Proof:
Let $x^1, ..., x^k$ be the vertices of P and $z^* := min\{c^T x x^i : i = 1, ..., k\}$. Given any $y \in P$, we need to prove that $c^T y \geq z^*$. Indeed, $y \in P$ implies the existence of multipliers $\lambda_1, ..., \lambda_k \geq 0, \Sigma_{i=1}^k \lambda_i = 1$, such that $y = \Sigma_{i=1}^k \lambda_i x^i$. Hence we have $c^T y = c^T \Sigma_{i=1}^k \lambda_i x^i = \Sigma_{i=1}^k \lambda_i (c^T x^i) \geq \Sigma_{i=1}^k \lambda_i z^* = z^*$.

Definition 4.1.5

A collection of m linearly independent columns of A is said to be a basis of A. The x_j variables associated with the basic columns are called basic variables; the remaining variables are called non-basic variables.

Definition 4.1.6

The solution obtained imposing $x_F = 0$ and $x_b = B^{-1}b$ is said to be the basic solution associated with basis B. The basic solution (and by extension, basis B itself) is said to be feasible if $x_B = B^{-1}b \geq 0$.

Definition 4.1.7

A basis B is said to be degenerate if $B^{-1}b$ has one or more zero components.

Theorem 4.1.3

A point $x \in P$ is a vertex of the not empty polyhedron $P := \{x \geq 0 : Ax = b\}$ if and only if x is a basic feasible solution of the system $Ax = b$.
Proof:
Let us prove the implication x is a basic feasible solution then x is a vertex. Let $x = [x_1, ..., x_k, 0, ..., 0]^T$ be any basic feasible solution associated with basis B of A, where $k \geq 0$ is the number of non-zero components of x. It follows that columns $A_1, ..., A_k$ must be part of B, possibly together with other columns. Let us assume by contradiction that x is not a vertex. There exist thus $y = [y_1, ..., y_k, 0, ..., 0]^T \in P$ and $z = [z_1, ..., z_k, 0, ..., 0]^T \in P$ with $y \neq z$ such that $x = \lambda y + (1 - \lambda)z$ for any $\lambda \in (0, 1)$, which implies that $k \geq 1$. Note that both y and z must have the last components set to zero, otherwise their convex combination cannot give x. For the hypothesis, we then have $y \in P \Rightarrow Ay = b \Rightarrow A_1y_1 + ... + A_ky_k = b$ and $z \in P \Rightarrow Az = b \Rightarrow A_1z_1 + ... + A_kz_k = b$. By subtracting the second equation from the first we obtain $(y_1 - z_1)A_1 + ... + (y_k - z_k)A_k = \alpha_1A_1 + ... + \alpha_kA_k = 0$, where $\alpha_i = y_i - z_i, i = 1, ..., k$. Hence there exist $\alpha_1, ..., \alpha_k$ scalars not all zero such that $\Sigma_{i=1}^{k}\alpha_iA_i = 0$, thus columns $A_1, ..., A_k$ are linearly dependent and cannot be part of the basis B ($\Rightarrow$ contradiction).
We will now prove the implication x is a vertex then x is a basic feasible solution; the fact that the basic solution is also feasible obviously derives from the hypothesis that $x \in P$. Writing, as before, $x = [x_1, ..., x_k, 0, ..., 0]^T$ with $x_1, ..., x_k > 0$ and $k \geq 0$, we have that $x \in P \Rightarrow Ax = b \Rightarrow A_1x_1 + ... + A_kx_k = b$. From here two cases can occur:

1. Columns $A_1, ..., A_k$ are linearly dependent (or k = 0): by arbitrarily selecting other m-k linearly independent columns (which, as is well known, is always possible), we obtain basis $B = [A_1, ..., A_k, ...]$ whose basic associated solution is indeed x (which satisfies Ax=b and has non-basic components all equal to zero), thus concluding the proof.

2. columns $A_1, ..., A_k$ are linearly dependent: we will prove that this case cannot actually happen. Indeed, if the columns were linearly dependent, then there would exist $\alpha_1, ..., \alpha_k$ not all zero such that $\alpha_1A_1 + ... + \alpha_kA_k = 0$. With some math manipulation we can obtain, with $\epsilon > 0$, the following $(x_1 + \epsilon\alpha_1)A_1 + .. + (x_k + \epsilon\alpha_k)A_k = b$ and $(x_1 - \epsilon\alpha_1)A_1 + .. + (x_k - \epsilon\alpha_k)A_k = b$. By defining $y := [x_1 - \epsilon\alpha_1, ..., x_k - \epsilon\alpha_k, 0, ..., 0]^T$ and $z := [x_1 + \epsilon\alpha_1, ..., x_k + \epsilon\alpha_k, 0, ..., 0]^T$, we would have Ay=b and Az=b, while choosing a sufficient small ϵ we would have $y, z \geq 0$ and thus $u, z \in P, y \neq z$. But since by construction we have $x = \frac{1}{2}y + \frac{1}{2}z$, thus would mean that vertex x can be expressed as the strict convex combination of two distinct points of P($\Rightarrow$ contradiction).

Corollary 4.1.1

Every problem $min\{c^Tx : Ax = b, x \geq 0\}$ defined on a polytope $P = \{x \geq 0 : Ax = b\} \neq \emptyset$ has at least one optimal solution coinciding with a basic feasible solution.
Proof:
According to theorem 4.1.2, there always exists an optimal solution coinciding with a vertex of P and thus, according to theorem 4.1.3, with a basic feasible solution.