3. Structura mulțimii soluțiilor admisibile ale unei probleme de programare liniară

În această secțiune ne vom opri asupra principalelor proprietăți geometrice pe care le posedă mulțimea soluțiilor unui sistem de ecuații și inecuații liniare. Aceste proprietăți sunt determinante în înțelegerea mecanismului metodei simplex de rezolvare a programelor liniare.

Parcurgerea acestei secțiuni necesită câteva rudimente de calcul matricial și algebră liniară. Vectorii cu care se va opera vor fi subînțeleși, după caz, fie linii fie coloane. De regulă,scrierea în text a unui vector se va face în linie ca de exemplu $v = (a_1, a_2, ..., a_m)$; dacă este necesar ca v să fie considerat vector coloană se va folosi operatorul de transpunere: $v = (a_1, a_2, ..., a_m)^{\mathrm{T}}$.

3.1 Câteva elemente de analiză convexă liniară

Fiind date două puncte $x, y \in R^n$ mulțimea:

$$[x, y] = \{z = (1-\alpha)x + \alpha y,\ 0 \le \alpha \le 1\}$$

se numește *segment* (închis) cu extremitățile *x* și *y.* Se știe că în R^2 sau în R^3 acest concept se suprapune peste conceptul geometric uzual. Pentru $\alpha = 0$, respectiv $\alpha = 1$,avem $z \equiv x$, respectiv $z \equiv y$. Punctele $z = (1-\alpha)x + \alpha y$ corespunzătoare valorilor $\alpha \in (0,1)$ se numesc puncte *interioare* ale segmentului $[x,y]$. Pentru $\alpha = \frac{1}{2}$ găsim $z = \frac{1}{2}(x+y) \equiv$ mijlocul segmentului $[x,y]$.

O mulțime $X \subseteq R^n$ *se zice convexă dacă o dată cu două puncte conține și segmentul care le unește.*

Formal :

$$X \text{ convexă} \Leftrightarrow (\forall) x, y \in X, (\forall)\ \alpha \in [0,1], z = (1-\alpha)x + \alpha y \in X$$

Se verifică imediat că intersecția mai multor mulțimi convexe este o mulțime convexă.

Fie $a = (a_1, a_2, ..., a_n)$ un vector nenul şi b un scalar.Este uşor de văzut că mulţimea:

$$S = \left\{ x = (x_1, x_2, \dots, x_n)^T \middle| \, ax \le b \Leftrightarrow a_1x_1 + a_2x_2 + \dots + a_nx_n \le b \right\}$$

este convexă. Ea se numeşte *semispaţiu*, în timp ce mulţimea:

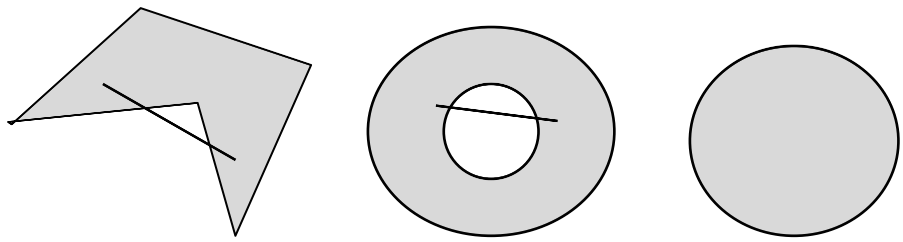
$$H = \left\{ x = (x_1, x_2, \dots, x_n)^T \middle| \, ax = b \Leftrightarrow a_1x_1 + a_2x_2 + \dots + a_nx_n = b \right\}$$

se numeşte *hiperplan*. Este clar că şi H este o mulţime convexă ca intersecţie a semispaţiului S de mai sus, cu semispaţiul:

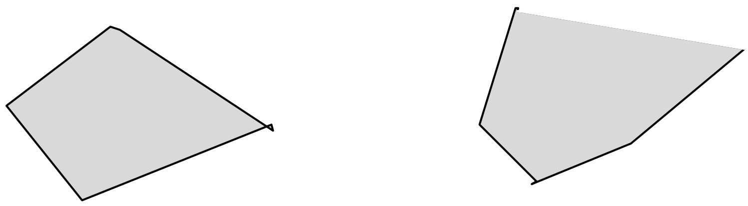
$$S' = \left\{ x \in R^n \middle| \, ax \ge b \Leftrightarrow (-a)x \le -b \Leftrightarrow -a_1x_1 - a_2x_2 - \dots - a_nx_n \le -b \right\}$$

O intersecţie finită de semispaţii se numeşte mulţime poliedrală. Evident, o mulţime poliedrală este convexă, reciproca nefiind în general adevărată.

În figura 3.1.1 sunt prezentate câteva mulţimi convexe şi neconvexe în plan. Este clar că mulţimile a) şi b) nu sunt convexe. Discul c) este o mulţime convexă dar nu este poliedrală, fiind în fapt intersecţia infinită a tuturor semispaţiilor care conţin discul şi sunt mărginite de tangentele la circumferinţă. Poligonul convex d) este intersecţia a 4 semispaţii aşa cum se arată în fig.3.1.2



a)poligon concav *b)coroană circulară* *c)disc*



d)poligon convex *e)mulţime poliedrală nemărginită*

Notă: Toate mulţimile specificate sunt presupuse, închise adică îşi conţin frontierele.

Figura 3.1.1

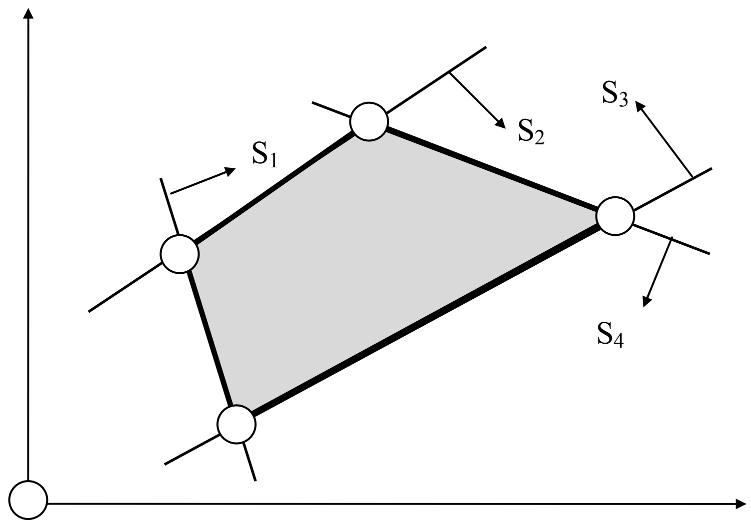


Figura 3.1.2

Din cele de mai sus rezultă că orice *mulțime poliedrală în* R^n *se identifică cu mulțimea soluțiilor unui sistem de ecuații şi/sau inecuații liniare* în *n* variabile. În particular:

Mulțimea $\mathcal{A}_P$ *a soluțiilor admisibile ale unui program liniar* (P) *este o mulțime convexă, poliedrală şi închisă. Frontiera sa se compune din toate punctele ale căror coordonate satisfac cu egalitate cel puțin una din restricții.*

Se numeşte *vârf* al unei mulțimi convexe $X \subseteq R^n$ un punct $v \in X$ cu proprietatea că nu există un segment $[x,y] \subset X$ care să conțină pe v ca punct interior. În R^2 sau R^3 regăsim conceptul geometric uzual.

O mulțime poliedrală are întotdeauna un număr finit de vârfuri (posibil nici unul); de exemplu, poligonul d) din fig. 3.1.1 are patru vârfuri în timp ce un semispațiu nu are vârfuri. Discul c) are o infinitate de vârfuri: orice punct de pe circumferință are această calitate.

Mulțimile d) şi e) din fig.3.1.1 sunt amândouă poliedrale dar e) este *nemărginită*. Pentru a caracteriza această proprietate avem nevoie de un nou concept.

Un vector $w \in R^n$ se numeşte *rază extremă* pentru mulțimea convexă $X \subseteq R^n$ dacă pentru orice $x \in X$ şi $\alpha \geq 0$ rezultă $x + \alpha \cdot w \in X$,iar w nu poate fi scris sub forma $w=w^1+w^2$,w^1 şi w^2 având aceeaşi proprietate ca şi w.

De exemplu mulțimea poliedrală din fig. 3.1.3 are două raze extreme w^1 şi w^2. Suporturile acestor vectori sunt paralele cu cele două muchii dealungul cărora "se poate merge către infinit".

De reținut că *o mulține convexă este nemărginită dacă şi numai dacă are cel puțin o rază extremă iar o mulțime poliedrală are un număr finit de raze extreme.*

Să considerăm acum o mulțime oarecare L de puncte din R^n , finită sau infinită. O *combinație convexă* de puncte din L este un punct (vector) de forma:

$$x = \alpha_1 v^1 + \alpha_2 v^2 + ... + \alpha_p v^p \quad \text{cu } \alpha_1, \alpha_2, ..., \alpha_p \geq 0 \text{ şi } \alpha_1 + \alpha_2 + ... + \alpha_p = 1$$

unde $v^1, v^2, ..., v^p$ sunt puncte din L arbitrar alese, numărul punctelor fiind deasemeni variabil.

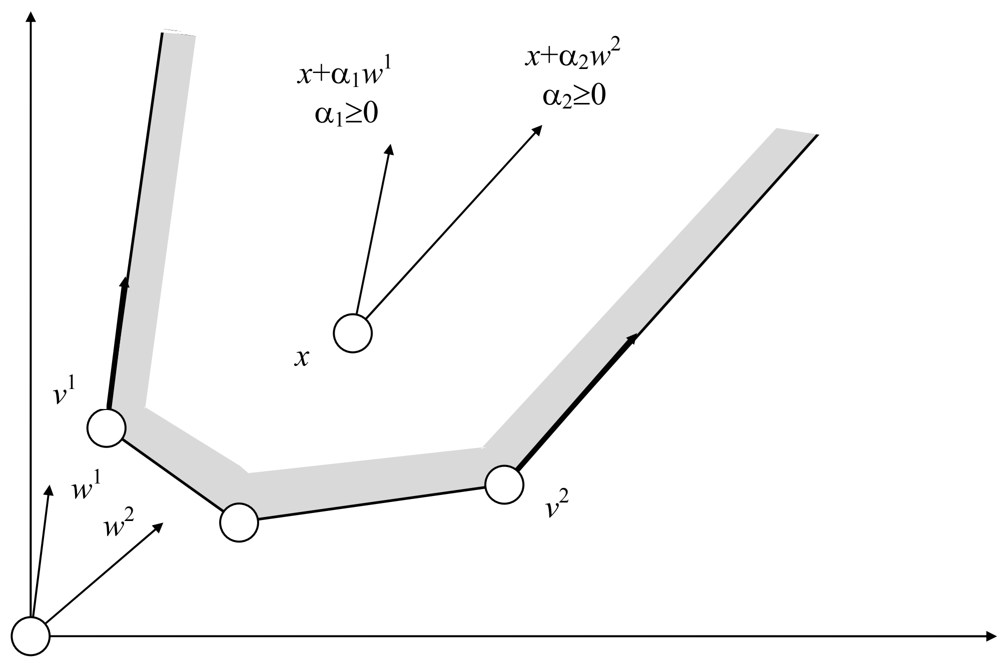


Figura 3.1.3

Se poate arăta fără dificultate că *mulțimea tuturor combinațiilor convexe de puncte din L este o mulțime convexă şi mai precis este cea mai mică mulțime convexă (în sensul incluziunii) care conține mulțimea* L.Ea se notează *conv L* şi se numeşte *anvelopa* (sau *acoperirea*) *convexă* a mulțimii L.

Dacă L este o mulțime finită atunci conv L este o mulțime poliedrală mărginită. Acest fapt este ilustrat în figura 3.1.4.

Să considerăm o mulțime poliedrală $X \subseteq R^n$ .Fie:

$$\dot{X} = \left\{v^1, v^2, \cdots, v^p\right\}$$

mulțimea vârfurilor sale. Deoarece $\dot{X} \subset X$ iar X este convexă avem $conv\,\dot{X} \subseteq X$. Se poate demonstra următorul rezultat clasic al analizei convexe:

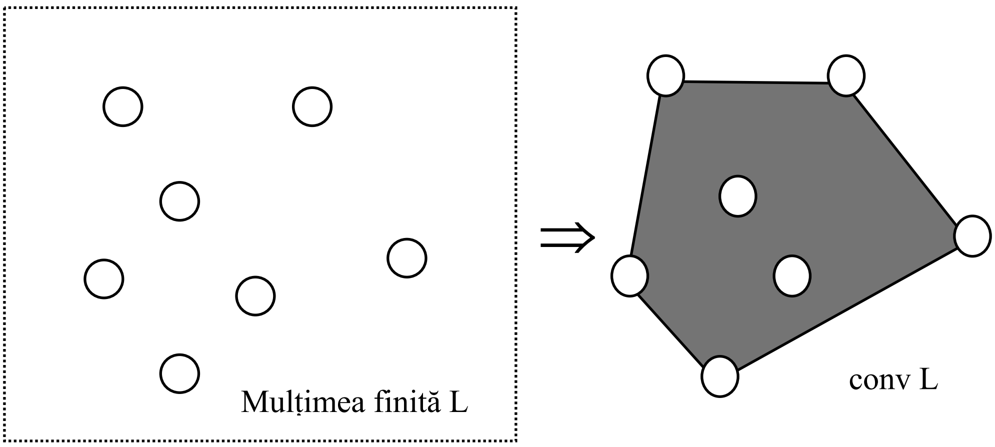


Figura 3.1.4

Dacă X este o mulțime poliedrală mărginită atunci $conv\,\dot{X} = X$, *adică orice punct din X este o combinație convexă a vârfurilor mulțimii X.*

Formal:

$$(\forall)x \in X, (\exists)\alpha_1, \alpha_2, \cdots, \alpha_p \geq 0 \text{ cu } \alpha_1 + \alpha_2 + \cdots + \alpha_p = 1 \text{ astfel încât :}$$

$$x = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_p v^p$$

De reținut faptul că scalarii $\alpha_1, \alpha_2, \cdots, \alpha_p$ nu sunt unici.

Dacă mulțimea poliedrală X nu este mărginită atunci $conv\dot{X} \neq X$. Acest fapt este ilustrat în figura 3.1.5 unde mulțimea dublu haşurată este acoperirea convexă a vârfurilor mulțimii poliedrale nemărginite X. Se poate arăta că dacă $v^1, v^2, \cdots, v^p$ sunt vârfurile mulțimii poliedrale nemărginite X iar $w^1, w^2, \cdots, w^q$ sunt razele sale extreme atunci pentru orice $x \in X$ există scalarii $\alpha_1, \alpha_2, \cdots, \alpha_p \geq 0$ cu $\alpha_1 + \alpha_2 + \cdots + \alpha_p = 1$ şi $\mu_1, \mu_2, \cdots, \mu_q \geq 0$ astfel încât:

$$x = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_p v^p + \mu_1 w^1 + \mu_2 w^2 + \cdots + \mu_q w^q$$

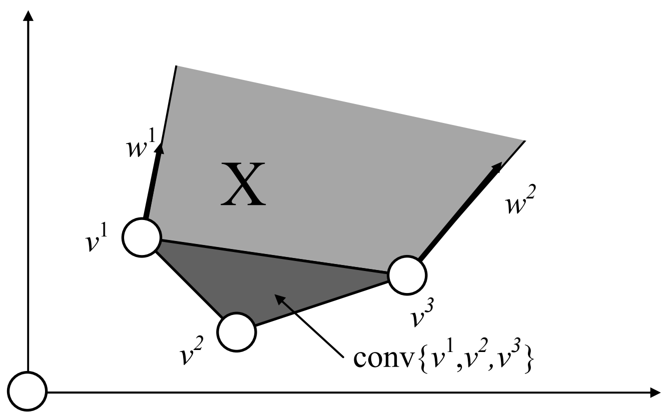


Figura 3.1.5

3.2 Teorema centrală a programării liniare

Să considerăm acum un program liniar (P) în care funcția obiectiv se maximizează şi să ne situăm în cazul în care programul (P) este compatibil adică mulțimea soluțiilor sale admisibile $\mathcal{A}_P$ este nevidă. Am văzut că $\mathcal{A}_P$ este o mulțime poliedrală şi convexă având un număr finit de vârfuri $v^1, v^2, \ldots\ldots, v^p$ şi (posibil) un număr finit de raze extreme $w^1, w^2, \cdots, w^q$. Aşa cum va rezulta din secțiunea 4.1, $\mathcal{A}_P$ are cel puțin un vârf, adică $p \geq 1$. Vom enunța şi demonstra acum *teorema centrală a programării liniare*.

TEOREMA . *Dacă programul (P) are optim finit, atunci o soluție optimă se găseşte într-unul din vârfurile mulțimii soluțiilor admisibile* $\mathcal{A}_P$.

Demonstrație: Fie:

$$f(x) = cx = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

funcția obiectiv a programului (P). Pentru început să arătăm că:

$$cw^k \leq 0 \quad (\forall)\ k = 1,\cdots,q$$

Să presupunem prin absurd că există o rază extremă $w \in \{w^1, w^2, \cdots, w^q\}$ astfel că $cw > 0$. Luăm o soluție admisibilă oarecare x^0 ($A_P \neq \emptyset$!) .Ştim că:

$$x(\alpha) = x^0 + \alpha w \in A_P\ ,\ (\forall)\ \alpha \geq 0$$

Atunci:

$$f(x(\alpha)) = cx^0 + \alpha cw \to \infty \quad \text{cînd} \quad \alpha \to \infty\,.$$

Ar rezulta că (P) are optim infinit contrar ipotezei.

Fie acum $v^* \in \{v^1, v^2, \cdots, v^p\}$ cu proprietatea că:

$$f(v^*) = cv^* = \max\left\{f(v^1) = cv^1, f(v^2) = cv^2, \cdots, f(v^p) = cv^p\right\}$$

Să considerăm o soluție admisibilă oarecare $x \in A_P$. Am văzut că există scalarii:

$\alpha_1, \alpha_2, \cdots, \alpha_p \geq 0$ cu $\alpha_1 + \alpha_2 + \cdots + \alpha_p = 1$ şi $\mu_1, \mu_2, ..., \mu_q \geq 0$ astfel încât:

$$x = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_p v^p + \mu_1 w^1 + \mu_2 w^2 + \cdots + \mu_q w^q$$

Atunci:

$$f(x) = cx = \alpha_1 cv^1 + \alpha_2 cv^2 + \ldots + \alpha_p cv^p + \mu_1 cw^1 + \mu_2 cw^2 + \ldots + \mu_q cw^q \leq$$
$$\leq \alpha_1 cv^1 + \alpha_2 cv^2 + \ldots + \alpha_p cv^p \leq (\alpha_1 + \alpha_2 + \ldots + \alpha_p) cv^* = f(v^*)$$

În consecință v^* este o soluție optimă a programului (P). ■

Importanța acestei teoreme este covârşitoare: ea reduce problema găsirii unei soluții optime x^* din mulțimea, în general infinită, $\mathcal{A}_P$ a tuturor soluțiilor admisibile ale programului (P), *la identificarea acestei soluții în mulțimea finită a vârfurilor lui $\mathcal{A}_P$.*

Recapitulând modul în care diferitele proprietăți discutate au fost implicate în obținerea acestui rezultat fundamental să reținem că:

- *Convexitatea mulțimii soluțiilor admisibile $\mathcal{A}_P$ situează soluțiile optime, dacă acestea există, pe frontiera lui $\mathcal{A}_P$;*
- *Deoarece $\mathcal{A}_P$ este poliedrală, iar funcția obiectiv este liniară, cel puțin una din soluțiile optime este un vârf al lui $\mathcal{A}_P$.*

Teorema furnizează următorul procedeu "naiv" de rezolvare a unui program liniar (P):

- se "generează" lista (finită) a vârfurilor mulțimii $\mathcal{A}_P$;
- prin înlocuire succesivă în funcția obiectiv se reține vârful care oferă acesteia valoarea maximă.

Procedeul ridică la rândul său următoarele probleme principiale:

1) Cum recunoaştem compatibilitatea programului (P) ?
2) Cum "calculăm" un vârf al mulțimii $\mathcal{A}_P$ sau mai corect spus cum se caracterizează "algebric" un vârf ?
3) Pentru obținerea soluției optime este necesar să generăm toate vârfurile mulțimii $\mathcal{A}_P$? Întrebarea este serioasă deoarece şi pentru programe liniare de dimensiuni reduse (adică cu un număr relativ mic de restricții şi variabile) numărul vârfurilor este foarte mare.
4) Chiar dacă reuşim, prin enumerarea explicită a tuturor vârfurilor, să găsim pe acela care maximizează funcția obiectiv, aceasta nu înseamnă obligatoriu că am rezolvat programul dat ! Este posibil ca programul respectiv să aibe optim infinit! Cum se recunoaşte acest fapt?

Vom răspunde progresiv la toate chestiunile menționate în secțiunile următoare.

3.3 Corespondența $\mathcal{A}_P \sim \mathcal{A}_{FSP}$

Considerăm o problemă de programare liniară (P) care conține cel puțin o restricție inegalitate şi fie (FSP) forma sa standard.(vezi secțiunea 1.4) Pentru simplificarea notațiilor, vom presupune că (P) este în formă canonică de maximizare:

$$(P)\begin{cases} \max f = cx \\ Ax \le b \\ x \ge 0 \end{cases}$$

unde:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad (3.3.1)$$

$$c = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$$

Ştim din secțiunea 1.3 că orice program liniar poate fi scris în această formă. Forma standard a programului (P) va fi:

$$(FSP)\begin{cases} max\ f = cx \\ Ax + y = b \\ x \ge 0\,,\ y \ge 0 \end{cases} \text{ în care: } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \text{ este vectorul variabilelor de abatere.}$$

Între mulțimile de soluții admisibile $\mathcal{A}_P \subset R^n$ şi $\mathcal{A}_{FSP} \subset R^{n+m}$ există o *corespondență bijectivă* $\Phi(x) = (x\ ,\ y)$, unde $y = b - Ax$, a cărei inversă este proiecția $\Phi^{-1}(x,y) = x$. Am remarcat deja în secțiunea 1.4 că prin corespondența

Φ, soluțiile optime ale celor două probleme se corespund. În fapt, Φ are următoarea proprietate mai generală.

Teorema 3.3.1 *Dacă x este un vârf al mulțimii* $\mathcal{A}_P$ *atunci* $\Phi(x) = (x,y)$ *cu* $y = b - Ax$ *este un vârf al mulțimii* $\mathcal{A}_{FSP}$ *. Reciproc, dacă* (x,y) *este un vârf al mulțimii* $\mathcal{A}_{FSP}$ *atunci* $\Phi^{1}(x,y) = x$ *este un vârf al mulțimii* $\mathcal{A}_P$ *.*

Demonstrație: Concluzia decurge nemijlocit din definiția vârfului (secțiunea 3.1) observând că Φ şi Φ^{-1} conservă combinațiile convexe ale punctelor din $\mathcal{A}_P$ şi $\mathcal{A}_{FSP}$.
Într-adevăr, fie x^1 , $x^2 \in \mathcal{A}_P$,$\alpha \in [0 , 1]$ şi $x = (1-\alpha)x^1 + \alpha x^2$. Fie mai departe: $\Phi(x) = (x,y)$, $\Phi(x^1) = (x^1,y^1)$, $\Phi(x^2) = (x^2,y^2)$ unde: $y = b - Ax$, $y^1=b - Ax^1$, $y^2 = b - Ax^2$ .Deoarece:
$y = b - A((1 - \alpha)x^1 + \alpha x^2)= (1 - \alpha)(b - A x^1) + \alpha (b - A x^2) = (1 - \alpha)y^1 + \alpha y^2$
avem:
$(x,y) = (1 - \alpha)(x^1 ,y^1) + \alpha (x^2 ,y^2),$ adică $\Phi(x) = (1 - \alpha)\Phi(x^1) + \alpha \Phi(x^2)$. ■

În baza acestei teoreme precum şi a teoremei centrale a programării liniare (secțiunea 3.2), pentru a rezolva problema (P) este suficient să căutăm soluția optimă a formei sale standard (FSP) printre vârfurile mulțimii $\mathcal{A}_{FSP}$.

Vom vedea în secțiunea următoare cum se caracterizează "algebric" vârfurile mulțimii soluțiilor admisibile ale unui program liniar în formă standard. Tot acolo vom arăta că dacă un program (P) este compatibil atunci $\mathcal{A}_P$ are cel puțin un vârf şi în orice caz un număr finit de asemenea elemente. Pe baza acestor rezultate, vom putea descrie în paragraful 4 o metodă efectivă de rezolvare a unei probleme de programare liniară.

3.4 Soluții de bază ale unui program liniar

Să considerăm acum un program liniar (P) în formă standard:

$$(P)\begin{cases} max\ f = cx \\ Ax = b \\ x \geq 0 \end{cases}$$

în care masivele A,b,c,x au semnificațiile din (3.3.1). Vom pune în evidență coloanele matricii A:

$$A = [A^1, A^2, \dots, A^n]$$

Definiție *Soluția* $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ *a problemei în formă standard (P), nu neapărat admisibilă, se numeşte soluție de bază dacă mulțimea coloanelor* A^i *corespunzătoare componentelor* $\bar{x}_i \neq 0$ *este liniar independentă.*

Fie I($\bar{x}$) mulțimea indicilor i ∈ {1,2,…,m} cu proprietatea că $\bar{x}_i \neq 0$.

Lema 1 *Dacă* $\bar{x}$ *şi* $\bar{x}'$ *sunt soluții de bază ale programului (P) şi* $I(\bar{x}') \subseteq I(\bar{x})$ *atunci* $\bar{x} = \bar{x}'$ *(şi deci* $I(\bar{x}) = I(\bar{x}')$*).*

Demonstrație: Este clar că $\bar{x}_i = \bar{x}_i' = 0$ pentru indicii $i \notin I(\bar{x})$. Atunci egalitățile:

$$\sum_{i \in I(\bar{x})} \bar{x}_i A^i = \sum_{i \in I(\bar{x}')} \bar{x}_i' A^i = b$$

implică:

$$\sum_{i \in I(\bar{x})} (\bar{x}_i - \bar{x}_i') A^i = 0$$

de unde rezultă $\bar{x}_i = \bar{x}_i'$ pentru toți $i \in I(\bar{x})$, deoarece prin ipoteză coloanele A^i, $i \in I(\bar{x})$ sunt liniar independente. ■

Lema 2 *Fie* $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ *o soluție admisibilă a problemei (P) care nu este soluție de bază. Atunci există un vector* $y \in R^n$ *şi un interval* $[\underline{\lambda}, \bar{\lambda}] \subset R^n \cup \{-\infty, +\infty\}$ *astfel încât:*

1) $Ay = 0$;

2) $[\underline{\lambda}, \bar{\lambda}]$ *conține pe 0 şi nu se reduce la acest punct;* $\underline{\lambda}$ *şi* $\bar{\lambda}$ *nu sunt simultan infinite;*

3) pentru orice $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ *vectorul* $x(\lambda) = \bar{x} + \lambda y$ *este o soluție admisibilă a problemei (P);*

4) Dacă de exemplu, $\bar{\lambda}$ *este finit şi* $\bar{x}' = x(\bar{\lambda})$ *atunci* $I(\bar{x}') \subseteq I(\bar{x})$ *dar* $|I(\bar{x}')| \leq |I(\bar{x})| - 1$ *adică* $\bar{x}'$ *are mai puține componente nenule decât* $\bar{x}$.

Demonstrație: Din ipoteză rezultă că mulțimea coloanelor A^i, $i \in I(\bar{x})$ este liniar independentă. Există prin urmare scalarii y_i, $i \in I(\bar{x})$ nu toți nuli astfel încât:

$$\sum_{i \in I(\bar{x})} y_i A^i = 0$$

Punând $y_i = 0$ pentru $i \notin I(\bar{x})$ obținem un vector $y = (y_1, y_2, \dots, y_n) \in R^n$ cu proprietatea că:

$$\sum_{i=1}^{n} y_i A^i = 0 \quad \Leftrightarrow \quad Ay = 0$$

Afirmația 1) este demonstrată.

Pentru orice $\lambda \in R$ vectorul $x(\lambda) = \bar{x} + \lambda y$ este o soluție a problemei (P) deoarece $Ax(\lambda) = A\bar{x} + \lambda Ay = b$.
Impunând condiția de admisibilitate $x(\lambda) \geq 0$ obținem pentru λ intervalul de valori permise $[\ \underline{\lambda}\ ,\ \overline{\lambda}\ ]$ în care:

$$\underline{\lambda} = \begin{cases} \max\limits_{i \in I(\bar{x}), y_i > 0} \left\{ -\bar{x}_i / y_i \right\} \\ -\infty \ \text{ daca toti } \ y_i \leq 0 \end{cases} \qquad \overline{\lambda} = \begin{cases} \min\limits_{i \in I(\bar{x}), y_i < 0} \left\{ -\bar{x}_i / y_i \right\} \\ +\infty \ \text{ daca toti } \ y_i \geq 0 \end{cases}$$

Avem $\underline{\lambda} < 0 < \overline{\lambda}$ și deoarece $y \neq 0$, cel puțin una din extremitățile $\underline{\lambda}, \overline{\lambda}$ este finită. Astfel și afirmațiile 2) , 3) sunt probate.

Să presupunem, în final că $\overline{\lambda}$ este finit. Atunci va exista un indice $r \in I(\bar{x})$ astfel că: $y_r < 0$ și $\overline{\lambda} = -\bar{x}_r / y_r$. Dacă $\bar{x}' = x(\overline{\lambda})$ este clar că $I(\bar{x}') \subseteq I(\bar{x})$ și cum $\bar{x}'_r = \bar{x}_r + \overline{\lambda} y_r = 0$ iar $\bar{x}_r > 0$ urmează că $|I(\bar{x}')| < |I(\bar{x})|$ și ultima afirmație este dovedită. ■

Teorema 3.4.1 *O soluție admisibilă* $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ *a problemei (P) este un vârf al mulțimii* $\mathcal{A}_P$ *dacă și numai dacă* $\bar{x}$ *este o soluție de bază.*

Demonstrație: Să presupunem că $\bar{x}$ este vârf dar nu este soluție de bază. Conform lemei 2 există $y \in R^n$ cu $Ay = 0$ și intervalul $[\ \underline{\lambda}\ ,\ \overline{\lambda}\ ]$ care conține pe 0 și nu se reduce la acesta, astfel încât $x(\lambda) = \bar{x} + \lambda y$ să fie o soluție a programului (P) oricare ar fi $\lambda \in [\underline{\lambda}, \bar{\lambda}]$. Alegem $\varepsilon > 0$ suficient de mic astfel încât $[-\varepsilon, +\varepsilon] \subset [\underline{\lambda}, \bar{\lambda}]$ şi punem: $x^1 = \bar{x} - \varepsilon y$, $x^2 = \bar{x} + \varepsilon y$. Atunci $x^1, x^2 \in \mathcal{A}_P$, $x^1 \neq x^2$ şi $\bar{x} = \frac{1}{2}(x^1 + x^2)$ în contradicție cu ipoteza că $\bar{x}$ este vârf al mulțimii $\mathcal{A}_P$.

Pentru reciprocă să presupunem că $\bar{x}$ este o soluție de bază fără a fi vârf. Atunci există $x^1, x^2 \in A_P$, $x^1 \neq x^2$ şi $\alpha \in (0,1)$ astfel încât $\bar{x} = (1-\alpha)x^1 + \alpha x^2$. Pentru $i \notin I(\bar{x})$ avem $(1-\alpha)x_i^1 + \alpha x_i^2 = 0$ şi cum $x_i^1 \geq 0$, $x_i^2 \geq 0$ rezultă $x_i^1 = x_i^2 = 0$. În consecință, $I(x^1) \subseteq I(\bar{x})$, $I(x^2) \subseteq I(\bar{x})$ şi în virtutea lemei 1 rezultă $x^1 = x^2 = \bar{x}$, în contradicție cu ipoteza făcută. ■

Teorema 3.4.2 *Dacă programul în formă standard (P) este compatibil atunci mulțimea soluțiilor sale admisibile A_P are cel puțin o soluție de bază, deci un vârf.*

Demonstrație Fie $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)^T$ o soluție admisibilă a problemei (P). Vom proceda prin inducție după numărul *k* al componentelor $\bar{x}_i > 0$.Dacă k = 0 atunci $\bar{x}$ = (0,0,...,0) este o soluție de bază întrucât o mulțime vidă de vectori este, prin convenție,liniar independentă. Dacă k > 0 există două situații de examinat:

1) Coloanele A^i, $i \in I(\bar{x})$ sunt liniar independente. Atunci $\bar{x}$ este o soluție de bază.

2) Coloanele A^i, $i \in I(\bar{x})$ sunt liniar dependente. Conform lemei 2 va exista o soluție admisibilă $\bar{x}'$ cu $I(\bar{x}') \subset I(\bar{x})$ dar cu mai puține componente nenule decât $\bar{x}$. Repetând raționamentul, este clar că într-un număr finit de paşi se ajunge la situația 1) adică la o soluție de bază. ■

Consecință *Mulțimea $\mathcal{A}_P$ a soluțiilor admisibile ale unui program liniar compatibil are cel puțin un vârf.*

Demonstrație: Să presupunem că (P) nu este în formă standard, altminteri avem teorema 3.4.2. Fie (FSP) forma standard a programului (P). Deoarece avem corespondența bijectivă $\Phi : \mathcal{A}_P \cong \mathcal{A}_{FSP}$ deducem că $\mathcal{A}_{FSP} \neq \emptyset$ pentru că prin ipoteză $\mathcal{A}_P \neq \emptyset$. Prin teorema 3.4.2 mulțimea $\mathcal{A}_{FSP}$ are cel puțin un vârf; acesta, prin proiecția Φ^{-1} va fi un vârf al mulțimii $\mathcal{A}_P$, grație teoremei 3.3.1. ■

Având în vedere caracterizarea algebrică a vârfurilor dată în teorema 3.4.1, teorema centrală a programării liniare poate fi formulată şi în următorii termeni.

Teorema 3.4.3 *Dacă un program liniar în formă standard are optim finit, cel puţin una din soluţiile sale optime este o soluţie de bază.*

Mai rămâne de arătat că mulţimea soluţiilor admisibile ale unui program liniar are un număr finit de vârfuri. În virtutea teoremelor 3.3.1 şi 3.4.1, aceasta revine la a arăta că un program liniar în formă standard (P) are un număr finit de soluţii admisibile de bază. Faptul rezultă nemijlocit din aceea că numărul sistemelor liniar independente ce pot fi extrase dintr-o mulţime finită de vectori este finit. Vom preciza acest lucru sub forma unei teoreme în secţiunea următoare, în care vom introduce un concept uşor diferit de cel de soluţie de bază, acela de *soluţie asociată unei baze* a programului (P). Noul concept are avantajul de a fi mai uşor de manipulat în practică.

3.5 Baze ale unui program liniar în formă standard. Soluţia asociată unei baze

În notaţiile secţiunii precedente facem ipoteza:

$$\text{rang}A = m < n \tag{3.5.1}$$

Ipoteza ne asigură că ecuaţiile ce compun sistemul liniar $Ax = b$ al restricţiilor sunt independente şi că acest sistem are o infinitate de soluţii. Să notăm că ipoteza nu implică în mod necesar şi existenţa soluţiilor admisibile (adică cu toate componentele nenegative) pentru sistemul considerat. După cum vom vedea în secţiunea 4.6 ipoteza făcută nu este deloc restrictivă.

Deoarece rangA = m, în matricea A va exista cel puţin un grup de m coloane liniar independente, constituind deci o bază a spaţiului R^m. Un asemenea grup se va numi *bază a programului (P)* şi numărul acestora este finit , nedepăşind $C_n^m = \dfrac{n!}{m!(n-m)!}$.

Fie B o bază a programului (P), I mulţimea indicilor coloanelor din B şi J mulţimea indicilor coloanelor din A care nu sunt în B. Renumerotând convenabil variabilele programului vom scrie matricea A în forma:

$$A = [B , S] \quad \text{cu } B = [A^i]_{i \in I} \quad , \quad S = [A^j]_{j \in J}$$

Partiționăm corespunzător vectorul (coloană) x al variabilelor:

$$x = \begin{bmatrix} x^B \\ x^S \end{bmatrix} \qquad \text{cu } x^B = [x_i]_{i \in I} \quad , \quad x^S = [x_j]_{j \in J}$$

ca şi vectorul (linie) al coeficienților funcției obiectiv:

$$c = [c^B, c^S] \qquad \text{cu } c^B = [c_i]_{i \in I} \quad , \qquad c^S = [c_j]_{j \in J}$$

În raport cu baza B aleasă, variabilele x_i , $i \in I$ se vor numi *variabile bazice*, iar toate celelalte *nebazice* sau *secundare*.

Scriem sistemul Ax = b în forma:

$$[B,S]\begin{bmatrix} x^B \\ x^S \end{bmatrix} = b \quad \Leftrightarrow \quad Bx^B + Sx^S = b$$

Deoarece B este o bază a spațiului R^m, B (ca matrice!) este nesingulară deci inversabilă. Înmulțind la stânga cu B^{-1} obținem:

$$x^B + \overline{S}x^S = \overline{b} \quad \text{cu } \overline{S} = B^{-1}S = \left[\overline{a}_{ij}\right]_{i \in I, j \in J}, \overline{b} = B^{-1}b = \left[\overline{b}_i\right]_{i \in I} \qquad (3.5.2)$$

Va fi util să punem în evidență coloanele matricii $\overline{S}$:

$$\overline{S} = B^{-1}\left[A^j\right]_{j \in J} == \left[B^{-1}A^j\right]_{j \in J} = \left[\overline{A}^j\right]_{j \in J}$$

cu

$$\overline{A}^j = B^{-1}A^j = \left[\overline{a}_{ij}\right]_{i \in I} \qquad (3.5.3)$$

Utilizăm (3.5.2) pentru a elimina din expresia funcției obiectiv variabilele bazice:

$$f = \left[c^B, c^S\right]\begin{bmatrix} x^B \\ x^S \end{bmatrix} = c^B x^B + c^S x^S = c^B(\overline{b} - \overline{S}x^S) + c^S x^S =$$

$$= c^B\overline{b} - (c^B\overline{S} - c^S)x^S = \overline{f} - \overline{c}^S x^S \qquad (3.5.4)$$

unde:

$$f = c^B b\ (= c^B B^{-1} b) = \sum_{i \in I} c_i b_i \tag{3.5.5}$$

şi:

$$\overline{c}^S = c^B \overline{S} - c^S\ (= c^B B^{-1} S - c^S) = \left[\overline{c}_j\right]_{j \in J}$$

în care:

$$\overline{c}_j = c^B \overline{A}^j - c_j\ (= c^B B^{-1} A^j - c_j) = \sum_{i \in I} c_i \overline{a}_{ij} - c_j \tag{3.5.6}$$

Astfel, în raport cu baza B, programul (P) poate fi adus la următoarea formă echivalentă, conform (3.5.2) şi (3.5.4):

$$(P_B) \begin{cases} \max f = \bar{f} - \overline{c}^S x^S \\ x^B + \overline{S} x^S = \overline{b} \\ x^B \geq 0 \,, x^S \geq 0 \end{cases} \Leftrightarrow \begin{cases} \max f = \bar{f} - \sum_{j \in J} \overline{c}_j x_j \\ x_i + \sum_{j \in J} \overline{a}_{ij} x_j = \overline{b}_i \quad i \in I \\ x_i \geq 0 \,, i \in I \,; x_j \geq 0 \,, j \in J \end{cases} \tag{3.5.7}$$

(P_B) se numeşte *forma explicită a programului (P) în raport cu baza B*. Comparând (P_B) cu (P) constatăm că *efectul înmulţirii cu B^{-1} a sistemului $Ax = b$ este explicitarea variabilelor bazice x_i , $i \in I$ în funcţie de cele nebazice x_j , $j \in J$.*

Luând în (P_B):

$$x^S = 0 \quad \Leftrightarrow \quad x_j = 0 \,, j \in J \tag{3.5.8}$$

obţinem:

$$x^B = \overline{b} \quad \Leftrightarrow \quad x_i = \overline{b}_i \;\; i \in I \tag{3.5.9}$$

Vectorul:

$$\overline{x} = \begin{bmatrix} \overline{b} \\ 0 \end{bmatrix} \begin{matrix} \leftarrow x^B \\ \leftarrow x^S \end{matrix} \tag{3.5.10}$$

este evident o soluţie a programului (P), numită *soluţia asociată bazei B.* Vom spune că B este o *bază admisibilă* dacă soluţia asociată (3.5.10) este admisibilă, ceea ce revine la a spune că $\overline{b}_i \geq 0 \,, i \in I$.

Prin construcţie, *soluţia asociată unei baze a programului (P) este o soluţie de bază* în sensul definiţiei din secţiunea precedentă. Reciproca nu este în general adevărată. Astfel, dacă $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)^T$ este o soluţie de bază a programului (P), numărul componentelor nenule este ≤ *m*. Dacă acest număr este exact *m* atunci $\overline{x}$ este soluţia asociată bazei formate din coloanele matricii

A corespunzătoare celor m componente nenule. Spunem în acest caz că $\bar{x}$ este o *soluție de bază nedegenerată*. Dacă numărul componentelor nenule este $< m$,

coloanele corespunzătoare acestor componente pot fi completate până la o bază în mai multe moduri, astfel că $\bar{x}$ se asociază la mai multe baze! Dacă se întâmplă aşa spunem că $\bar{x}$ este o *soluție degenerată.*

Deoarece (P) are un număr finit de baze, el va avea şi un număr finit de soluții asociate acestor baze şi de aici un număr finit de soluții de bază.

În baza relațiilor (3.5.8), (3.5.9) constanta $\bar{f}$ din (3.5.5) reprezintă valoarea funcției obiectiv a programului (P) în soluția (3.5.10) asociată bazei B. Componentele $\bar{c}_j$, $j \in$ J (definite în (3.5.6)) ale vectorului $\bar{c}^S$ din (3.5.4) vor juca un rol esențial în caracterizarea optimalității unei soluții admisibile de bază. Ele se numesc *costuri reduse* şi sunt întotdeauna asociate variabilelor nebazice. Coeficienții numerici ai formei explicite (3.5.7) se trec într-un tabel (T_B) al cărui format este indicat în tabelul 3.5.1. Tabelul (T_B) se numeşte *tabelul simplex asociat bazei* B. Coloana "B" conține vectorii bazei B, coloana "CB" conține coeficienții din funcția obiectiv ai variabilelor bazice iar în coloana "VVB" apar valorile variabilelor bazice. În ultima linie a tabelului, numită şi *linia test*, apar valoarea $\bar{f}$ a funcției obiectiv în soluția asociată bazei B precum şi costurile reduse $\bar{c}_j$, $j \in$ J corespunzătoare coloanelor nebazice. Formula (3.5.5) arată că $\bar{f}$ este produsul scalar al coloanelor "CB" şi "VVB" în timp ce relația (3.5.6) arată că $\bar{c}_j$ se obține din produsul scalar al coloanelor "CB" şi "A^j" scăzând costul c_j scris deasupra coloanei "A^j".

				c_i		c_r		c_j		c_k	
CB	B	VVB	…	A^i	…	A^r	…	A^j	…	A^k	…
⋮	⋮	⋮		⋮		⋮		⋮		⋮	
c_i	A^i	$\bar{b}_i$	…	1	…	0	…	$\bar{a}_{ij}$	…	$\bar{a}_{ik}$	…
⋮	⋮	⋮		⋮		⋮		⋮		⋮	
c_r	A^r	$\bar{b}_r$	…	0	…	1	…	$\bar{a}_{rj}$	…	$\boxed{\bar{a}_{rk}}$	…
⋮	⋮	⋮		⋮		⋮		⋮		⋮	
	f	$\bar{f}$	…	*	…	*	…	$\bar{c}_j$	…	$\bar{c}_k$	…

(T_B)

Tabelul 3.5.1

Extindem definiția costului redus şi la coloanele "A^i", $i \in I$: $\overline{c}_i = c_i - c_i = 0$.Aceste costuri reduse au fost notate în linia test cu asteriscuri (*) spre a le deosebi de eventualele costuri reduse nule $\overline{c}_j$, $j \in J$ care, după cum vom vedea în secțiunea 4.2, indică - "la optim" - prezența mai multor soluții optime de bază.