PROGRAMACIÓ MATEMÀTICA

CARLOTA CORRALES LLAGOSTERA

Universitat Politècnica de Catalunya
Barcelona

ApuntsFME

Desembre 2017

TEMA 1: Introducció i fonaments

1.- Introducció

Definició. Problema de programació lineal.

Donats els vectors $c, l, u \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ i la matriu $A \in \mathbb{R}^{m \times n}$, es defineix el problema de programació lineal com:

$$(PL) \begin{cases} \min\limits_{x \in \mathbb{R}^n} z = c'x & \rightarrow \text{funció objectiu} \\ \text{s.a.:} & \\ \quad Ax \gtreqless b & \rightarrow \text{constriccions} \\ \quad l \leq x \leq u & \rightarrow \text{fites} \end{cases}$$

Definició. Regió factible de (PL)

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax \gtreqless b,\ l \leq x \leq u\}$$

És el conjunt de punts que satisfan totes les constriccions del problema.

Definició. Solució factible

$x \in \mathbb{R}^n$ factible de (PL) $\Leftrightarrow$ $x \in \mathcal{F}$

És qualsevol punt de la regió factible

Definició. Solució òptima

$x^* \in \mathcal{F}$ tq $c'x^* \leq c'y \quad \forall y \in \mathcal{F}$.

És el punt més allunyat de la direcció del gradient

⊛ Altres formes d'expressar (PL):

$$(PL) \quad \min_{x \in \mathbb{R}^n} \{c'x \mid x \in \mathcal{F}\}$$

$$(PL) = \begin{cases} \min\limits_{x \in \mathbb{R}^n} z = \sum\limits_{i=1}^{n} c_i x_i & \\ \text{s.a.:} & \\ \quad \sum\limits_{i=1}^{n} a_{ji} x_j \gtreqless b_j & j = 1, \dots, m \\ \quad l_i \leq x_i \leq u_i & i = 1, \dots, n \end{cases}$$

Exemple.

$$(PL) \begin{cases} \min z = x_1 + x_2 & \rightarrow \text{f. objectiu} \\ \text{s.a.:} & \\ \quad x_1 + x_2 \leq 1 & \\ \quad x_1 \geq 1/4 & \Big\} \rightarrow \text{constriccions} \\ \quad x_1, x_2 \geq 0 & \rightarrow \text{fites} \end{cases}$$

En aquest cas,

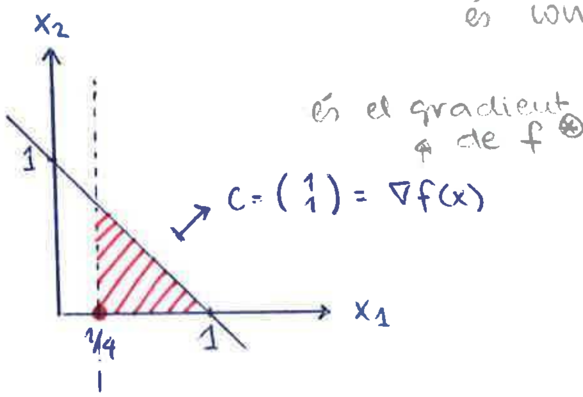
$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$l = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad u = \begin{pmatrix} +\infty \\ +\infty \end{pmatrix} \rightarrow \text{fites}$$

$$b = \begin{pmatrix} 1 \\ 1/4 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad Ax = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 \end{pmatrix} \leq b = \begin{pmatrix} 1 \\ 1/4 \end{pmatrix}$$

↓ la jacobiana és constant ⊛

Per a representar-ho:



/// = regió factible ($\mathcal{F}$)

$\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax \leq b, \ l \leq x \leq u\}$

$x \in \mathcal{F}$ = punt factible

• = punt òptim (solució òptima)

$x^* \in \mathcal{F} = \begin{pmatrix} 1/4 \\ 0 \end{pmatrix}$

⊛ Perquè estem treballant amb funcions lineals

2.- Models bàsics de programació lineal

PLANIFICACIÓ DE LA PRODUCCIÓ

↳ Manera òptima de fabricar una quantitat de producció per al qual utilitzem uns recursos comuns

⊛ Formulació parametritzada:

- Paràmetres: n = # productes a fabricar
  m = # recursos consumits
  c_i = benefici producte i; $i = 1, \ldots, n$
  b_j = disponibilitat recurs j; $j = 1, \ldots, m$
  a_{ji} = quantitat recurs j consumit per producte i
  l_i, u_i = fites a la producció
- Variables: x_i = quantitat a fabricar producte i
- Funció objectiu: es maximitzen els beneficis totals
- Constriccions: el programa de producció no consumeix més recursos dels existents.

(Hi ha un exemple al powerpoint)

PROBLEMA DE LA DIETA

↳ Determinar la manera més barata d'aconseguir tots els nutrients que he d'ingerir (sabent els nutrients que m'aporta cada aliment i el seu preu). L'objectiu és minimitzar el cost dels menjars.

⊛ Formulació parametritzada:

- Paràmetres: n = # aliments disponibles
  m = # nutrients essencials dieta
  c_i = cost aliment i ; $i = 1,\dots,n$
  b_j = quantitat diària mínima nutrient j ; $j = 1,\dots,m$
  a_{ji} = quantitat nutrient j per unitat aliment i
  l_i, u_i = fites quantitat element i.
- Variables: x_i = quantitat diària aliment i a la dieta
- Funció objectiu: es minimitza el preu total de la dieta
- Constriccions: la dieta aporta les quantitats necessàries de cada nutrient

(Hi ha un exemple al powerpoint)

PROBLEMA DE TRANSPORT

↳ tenim uns centres de producció i volem transportar la mercaderia als centres de consum de la manera òptima

⊛ Formulació parametritzada:

- Paràmetres: n = # centres producció
  m = # centres consum
  c_{ij} = cost unitari transport entre centres i,j
  p_i = producció centre i ; $i = 1,\dots,n$
  d_j = demanda centre consum j ; $j = 1,\dots,m$
- Variables: x_{ij} = quantitat de producte a transportar de i a j
- Funció objectiu: es minimitzen els costos totals de transport
- Constriccions: 1) Cada centre de producció envia tota la seva producció
  2) Cada centre de consum rep la seva demanda

⚠ La producció total ha de ser igual al consum total. (ha d'estar balancejat)
↓
equacions de balanç

PROBLEMES DE FLUXOS EN XARXES

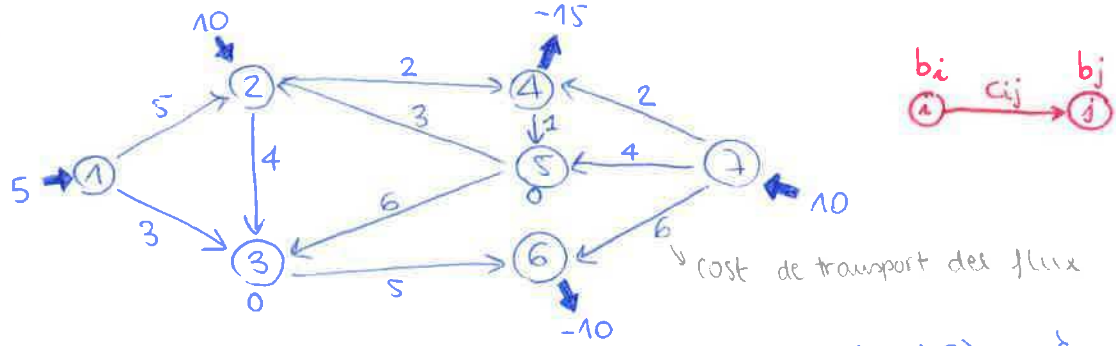
↳ Problemes d'optimització que es basen en una xarxa ⇒ hi associem un graf.

- Problemes de flux de cost mínim

↳ Trobar la forma més econòmica de transportar una mercaderia entre centres de producció i els centres de demanda a través d'una xarxa.

Sigui $G=(N,A)$ el graf dirigit definit per un conjunt N de n nodes i un conjunt A de m arcs dirigits.

Exemple.



$n=7$, $m=11$, $N=\{1,2,3,\dots,7\}$, $A=\{(1,2),(1,3),\dots\}$

- Vector de produccions/demandes: b_j, $j=1,\dots,n$: $b=[5,10,0,-15,0,5,10]'$
- Vector de costos: c_{ij}, $(i,j)\in A$: $c=[5,3,4,\dots]'$
- Fluxos: x_{ij}, $(i,j)\in A$: quantitat de mercaderia (flux) que circula per l'arc (i,j).

⊛ Formulació matemàtica:

$$
(PFCM)\left\{\begin{array}{l}
\min\limits_{x\in\mathbb{R}^{|A|}} \quad z=\sum\limits_{(i,j)\in A} c_{ij}x_{ij} \\
\text{s.a.:} \\
\qquad \underbrace{\sum\limits_{(i,j)\in A} x_{ij}}_{1)} - \underbrace{\sum\limits_{(j,i)\in A} x_{ji}}_{2)} = b_i \qquad i\in N \\
\qquad x_{ij}\geq 0 \qquad (i,j)\in A
\end{array}\right.
$$

- Les propietats d'aquesta matriu fan que es pugui accelerar molt la resolució del problema

1) Flux net de sortida
2) Flux net d'entrada

- Les primeres constriccions que veiem són les equacions de balanç
- La nostra hipòtesi és que $\sum_{i\in N} b_i = 0$ (xarxa balancejada)
- Propietat d'integritat dels fluxos en xarxes:

b_i enter, $i\in N$ ⇒ x_{ij}^* enter, $(i,j)\in A$

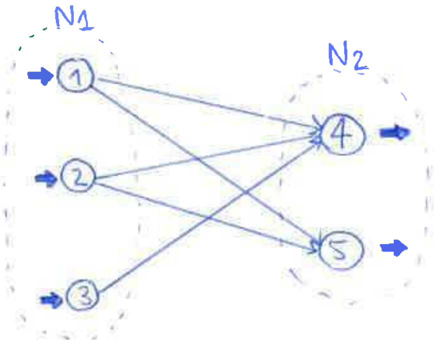
(En els PFCM la solució òptima és un nombre enter).

↳ Existeixen algorismes específics que saben aprofitar aquesta estructura

⊛ Problema de transport com a PFCM. És un cas particular dels PFCM, on tenim un graf bipartit. Tots els arcs dirigits van d'un centre de producció a un centre de consum.

- Característiques: PFCM sobre graf bipartit ($G=(N,A)$).
  - Tenim una partició de N ($N = N_1 \cup N_2$, $N_1 \cap N_2 = \emptyset$)
  - N_1 = nodes de producció, N_2 = nodes de demanda
  - $A \subseteq N_1 \times N_2$

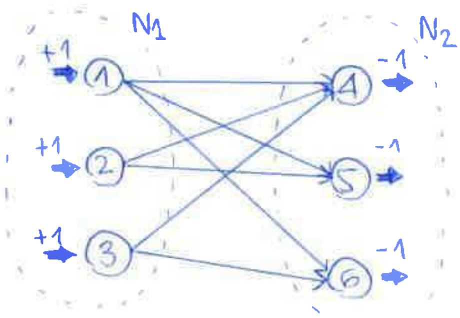
Exemple.



⊛ Problema d'assignació com a PFCM. Es tracta d'un cas particular del problema del transport com a PFCM. Per tant, podem utilitzar el mateix algorisme per a la seva resolució, tenint en compte que les variables de decisió (els fluxos x_{ij}) només poden prendre valors binaris (0 i 1) en la solució òptima.

- Característiques:
  - Tenim una partició de N i $|N_1| = |N_2|$ → (no sempre)
  - $A \subseteq N_1 \times N_2$
  - $\forall j \in N_1$, $b_j = +1$ i $\forall j \in N_2$, $b_j = -1$

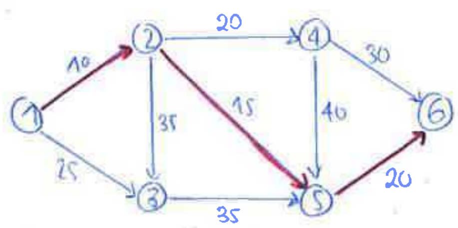
Exemple.



- Problemes de camins mínims

↳ L'objectiu és anar des d'un node de producció fins a un node de consum pel camí més curt

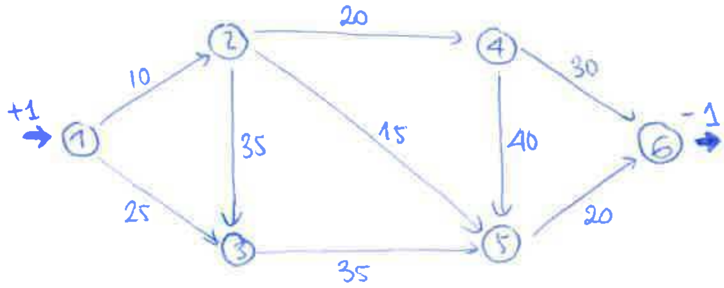
Exemple.



- El camí més curt per a anar de 1 a 6 seria: (1,2), (2,5), (5,6)

✻ <u>Problema de camins mínims com a PFCM</u>. És tracta d'un cas particular dels problemes de camins mínims, on volem transportar una unitat de flux des del node origen fins al node destí. Per tant en tots els altres nodes la quantitat de flux que entra és la mateixa que la que surt.

Exemple.

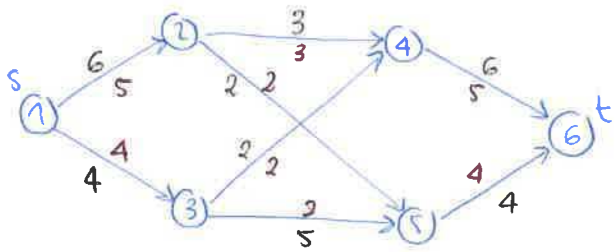


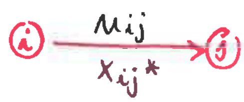
$$b_j = \begin{cases} +1 & j=1 \\ -1 & j=6 \\ 0 & j\notin\{1,6\} \end{cases}$$

- <u>Problemes de flux màxim</u>

↳ L'objectiu és trobar el <u>flux màxim</u> que es pot enviar entre un node origen i un node destí d'una xarxa <u>amb capacitats</u>

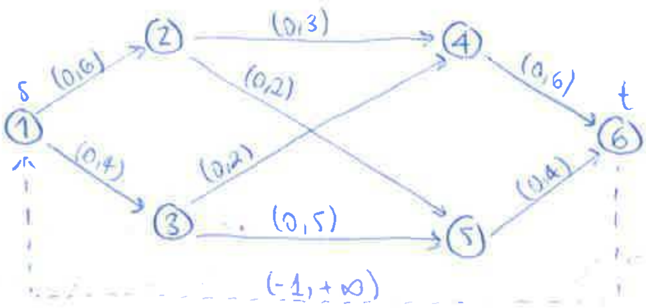
Exemple.

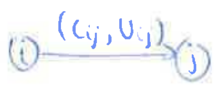




✻ <u>Problema de flux màxim com a PFCM</u>. És un cas particular dels problemes de flux màxim on l'objectiu és maximitzar la quantitat de flux que va des del node origen fins al node destí d'una xarxa amb capacitats. Equivalentment, podem crear una aresta imaginària amb capacitat infinita i maximitzar el flux que passarà per aquesta aresta.

Exemple.





✻ Transformem el problema en un problema de flux circulant

- Arc artificial x_{ts} amb $\underbrace{c_{ts}=-1}_{}$ i $u_{ts}=+\infty$
- $c_{ij}=0,\ (i,j)\neq(t,s)$

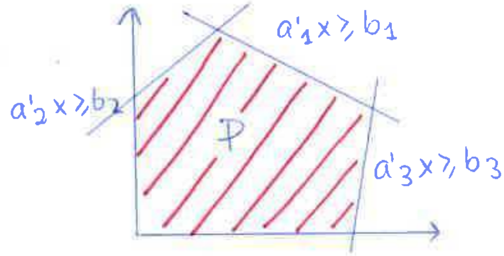
cost $=-1$ perquè és un problema de flux de cost mínim i així ens assegurem el flux de x_{ts}.

3.- Poliedres i politops

Definició. Poliedre

Un poliedre P és un conjunt de $\mathbb{R}^n$ que pot ser expressat com a intersecció d'una col·lecció finita de semiespais:

$P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

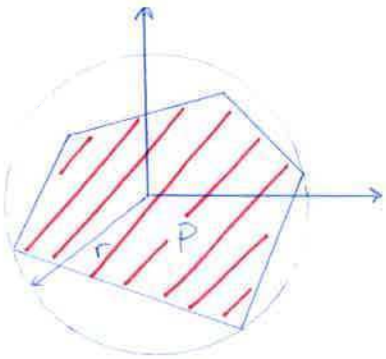


⊛ La regió factible de qualsevol problema (PL) és un poliedre:

$$(PL) \quad \min_{x \in \mathbb{R}^n} \{c'x \mid x \in \mathcal{F} \equiv P\}$$

Definició. Politop

Un politop és un poliedre no buit i fitat.



⊛ Els politops són conjunts compactes (tancats i fitats)

4.- Classificació dels problemes PL

Definició. Problema PL amb solució òptima

x^* solució òptima

És un problema PL tq $\exists x^* \in P \mid c'x^* \leq c'y \quad \forall y \in P,\ y \neq x^*$

(Es tracta d'un problema on, dins de la regió factible, hi ha una (o més d'una) solució que és "millor" que les altres perquè aconsegueix minimitzar el cost o maximitzar el benefici)

⊛ Dit d'una altra manera, és un problema PL tq $\{c'x \mid x \in P\}$ està fitat inferiorment.

Si x^* no és únic, les (infinites) solucions òptimes de PL s'anomenen òptims alternatius o bé conjunt òptim

Definició. Problema PL infactible.

És un problema PL on $P = \emptyset$ (entenem P com la regió factible)

poliedre

Definició. Problema PL il·limitat.

- Problema PL factible tq $\{c'x \mid x \in P\}$ no està fitat inferiorment
- Problema PL factible tq $\exists x \in P$, $d \in \mathbb{R}$ que satisfan:
  - i) $x + \varphi d \in P \quad \forall \varphi > 0$ (diem que d és un raig del poliedre P)
  - ii) $c'd < 0$ (diem que d és una direcció de descens sobre x).

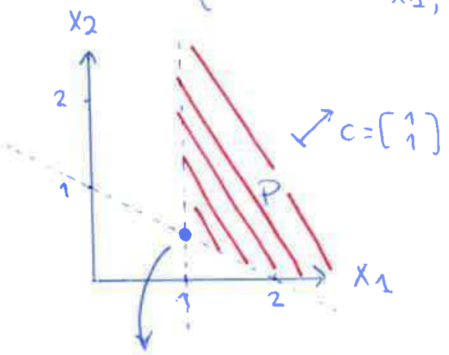
Exemples.

1) PL amb solució òptima

$$(PL) \begin{cases} \min\limits_{x \in \mathbb{R}^n} \ z = x_1 + x_2 \\ \text{s.a.:} \ x_1 + 2x_2 \geq 2 \\ \quad x_1 \geq 1 \\ \quad x_1, x_2 \geq 0 \end{cases}$$

$c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ $\quad b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$l = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad u = \begin{bmatrix} +\infty \\ +\infty \end{bmatrix}$

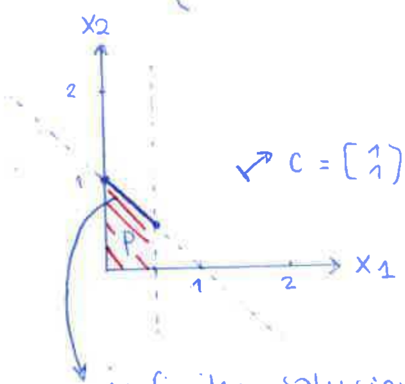


solució única: $x^* = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$ $\Rightarrow$ SOLUCIÓ ÒPTIMA

$$(PL) \begin{cases} \max\limits_{x \in \mathbb{R}^n} \ z = x_1 + x_2 \\ \text{s.a.:} \ x_1 + x_2 \leq 1 \\ \quad 2x_1 \leq 1 \\ \quad x_1, x_2 \geq 0 \end{cases}$$

$c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ $\quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$l = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad u = \begin{bmatrix} +\infty \\ +\infty \end{bmatrix}$



infinites solucions: $P^* = \{x \in \mathbb{R}^2 \mid \lambda x_1 + (1-\lambda)x_2, \lambda \in [0,1]\}$

$\Downarrow$

CONJUNT ÒPTIM

⚠ En aquest cas volíem MAXIMITZAR la funció!

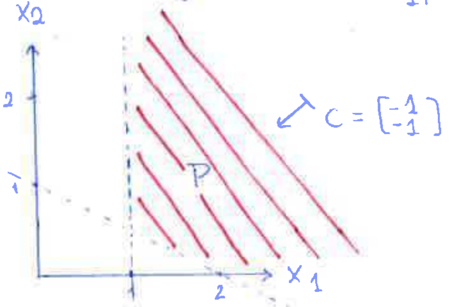
(si haguéssim volgut el mínim hi havia una única solució òptima: $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$)

2) PL il·limitat ($\nexists$ mínim)

$$(PL)\begin{cases} \min\limits_{x\in\mathbb{R}^n} & z = -x_1 - x_2 \\ s.a.: & x_1 + 2x_2 \geq 2 \\ & x_1 \geq 1 \\ & x_1, x_2 \geq 0 \end{cases}$$

$c = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ $b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$l = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $u = \begin{bmatrix} +\infty \\ +\infty \end{bmatrix}$



Per a trobar la solució òptima ens haurem de moure en la direcció ↗ (contrària al vector c), però la regió factible no està fitada per tant no existeix una solució òptima.

és un poliedre però NO és un politop

3) PL infactible ($\nexists$ solució factible)

$$(PL)\begin{cases} \max\limits_{x\in\mathbb{R}^n} & z = 2x_1 + x_2 \\ s.a.: & x_1 + x_2 \leq -1 \\ & 2x_1 \leq -1 \\ & x_1, x_2 \geq 0 \end{cases}$$

$c = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ $b = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

$l = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $u = \begin{bmatrix} +\infty \\ +\infty \end{bmatrix}$

No cal representar-ho gràficament per a veure que és incompatible que $2x_1 \leq -1$ i alhora $x_1 \geq 0$

Per tant, aquest problema PL no tindrà una regió factible ($P = \emptyset$) i en conseqüència no tindrà solució.

5.- Forma estàndard de P i PL.

Definició. Poliedre en forma estàndard

$P_e = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$.

Solució d'un sistema d'equacions lineals on les variables són ≥ 0.

Propietats:

i) P_e és un poliedre $Ax = b,\ x \geq 0 \Rightarrow \begin{bmatrix} A \\ -A \\ I \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix}$

ii) Tot poliedre $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ es pot expressar com a poliedre en forma estàndard

Definició. PL en forma estàndard

$$(PL)_e \quad \min_{x\in\mathbb{R}^n} z = \{c'x \mid x \in P_e\}$$

Proposició. Tot problema (PL) es pot transformar en un problema equivalent en forma estàndard $(PL)_e$, en el sentit que:

a) donada una solució factible d'un problema podem trobar una solució factible de l'altre amb el mateix cost

b) les solucions òptimes de (PL) i $(PL)_e$ coincideixen.

Regles pràctiques de pas a la forma estàndard

(PL)		(PL)e
• $a'_j x \leq b_j$	$\longrightarrow$	$a'_j x + x_K = b_j,\ x_K \geq 0$ (x_K variable de folga)
• $a'_j x \geq b_j$	$\longrightarrow$	$a'_j x - x_K = b_j,\ x_K \geq 0$ (x_K variable d'escreix)
• $\underline{b}_j \leq a'_j x \leq \bar{b}_j$	$\longrightarrow$	$a'_j x + x_K = \bar{b}_j,\quad 0 \leq x_K \leq \bar{b}_j - \underline{b}_j$
• $x_i \leq 0$	$\longrightarrow$	canvi de variable: $y_i = -x_i,\quad y_i \geq 0$
• x_i lliure	$\longrightarrow$	Mètode 1: $x_i = u_i - v_i,\ u_i, v_i \geq 0$ ($x_i = x_i^+ - x_i^-$) → una part positiva i una de negativa Mètode 2: s'elimina x_i d'una constricció $a'_j x = b_j$
• $x_i \leq u_i$	$\longrightarrow$	canvi de variable: $y_i = u_i - x_i,\quad y_i \geq 0$
• $l_i \leq x_i$	$\longrightarrow$	canvi de variable: $y_i = x_i - l_i,\quad y_i \geq 0$
• $\max c'x$	$\longrightarrow$	$\min -c'x$

Exemple. Volem transformar aquest problema PL a la forma estàndard:

$$(PL)\begin{cases} \min\limits_{x \in \mathbb{R}^n} z = 2x_1 - x_2 & \to \text{f. objectiu} \\ \text{s.a.: } 2x_1 - 3x_2 \geq 9 & \\ x_1 + 3x_2 = 2 & \text{constriccions} \\ 2x_1 - x_2 \leq -2 & \\ x_1 \leq 0 & \to \text{fites} \end{cases}$$

Hem d'aconseguir que ens quedi un problema PL de la forma $Ax = b$, $x \geq 0$.

La variable x_1 la tenim fitada $x_1 \leq 0$. Farem un canvi de variable tq $\tilde{x}_1 = -x_1$, per tant la fita ens quedarà $\tilde{x}_1 \geq 0$.

⚠ En les altres constriccions també hem de fer el canvi de variable.

- $-2\tilde{x}_1 - 3x_2 \geq 9$. Hem de posar-hi una variable d'escreix:

  $-2\tilde{x}_1 - 3x_2 - x_3 = 9,\quad x_3 \geq 0$ — x_3 és la variable d'escreix

- $-\tilde{x}_1 + 3x_2 = 2$ OK.

- $-2\tilde{x}_1 - x_2 \leq -2$. Hem de posar-hi una variable de folga

  $-2\tilde{x}_1 - x_2 + x_4 = 2,\quad x_4 \geq 0$ — x_4 és la variable de folga

Finalment, no tenim cap fita per a la variable x_2, se'n diu variable lliure. Utilitzarem el mètode 1: $x_2 = x_2^+ - x_2^-$; $x_2^+, x_2^- \geq 0$

Definició. Variable lliure

És aquella variable que pot adoptar valors tant positius com negatius.

Les constriccions quedaran $-2\tilde{x}_1 - 3x_2^+ + 3x_2^- - x_3 = 9$, etc
i la f. objectiu també queda amb les noves variables.

Exemple. Transformar el problema PL en forma estàndard

$$(PL)\begin{cases} \max\limits_{x\in\mathbb{R}^n} z = 3x_1+2x_2+3x_3 \\ \text{s.a.:}\quad 4\leq 2x_1+x_2+x_3\leq 20 \\ \qquad 3x_1-x_2+2x_3\leq 6 \\ \qquad x_1\geq 0\ ,\ x_2\geq 3 \end{cases}$$

$x_1 \geq 0$ ✓

$x_2 \geq 3 \Rightarrow \tilde{x}_2 = x_2 - 3\ ,\ \tilde{x}_2 \geq 0$ ✓

x_3 lliure $\Rightarrow x_3 = x_3^+ - x_3^-\ ;\ x_3^+, x_3^- \geq 0$ ✓

- $4 \leq 2x_1 + x_2 + x_3 \leq 20$

  $1 \leq 2x_1 + \tilde{x}_2 + x_3 \leq 20$

  $2x_1 + \tilde{x}_2 + x_3 + x_4 = 20\ ;\quad 0 \leq x_4 \leq 19$

$x_4 \leq 19 \Rightarrow \tilde{x}_4 = 19 - x_4\ ;\ \tilde{x}_4 \geq 0$ ✓

Per tant, $\boxed{2x_1 + \tilde{x}_2 + x_3 - \tilde{x}_4 = 1}^{(r1)}$

- $3x_1 - x_2 + 2x_3 \leq 6$

  $3x_1 - \tilde{x}_2 + 2x_3^+ - 2x_3^- \leq 9$

  $\boxed{3x_1 - \tilde{x}_2 + 2x_3^+ - 2x_3^- + x_5 = 9}^{(r2)},\ x_5 \geq 0$ ✓

⊛ Quan convertim el problema PL en estàndard, quasibé sempre acabem augmentant-ne la dimensió (En aquest últim exemple hem passat de 3 a 6 variables).

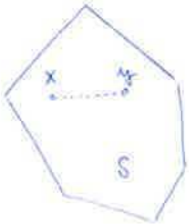
Però fer el canvi a forma estàndard simplifica la forma com es poden manipular els algorismes, per això ho fem.

6.- Convexitat i poliedres. Punts extrems

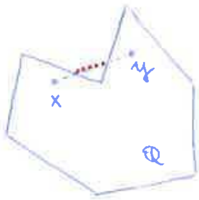
Definició. Conjunt convex.

Un conjunt $S \subset \mathbb{R}^n$ és convex si $\forall x, y \in S$, $\forall \lambda \in [0,1]$ $\lambda x + (1-\lambda)y \in S$.

(És a dir, si agafo 2 punts qualssevol del conjunt i els uneixo amb una recta, el segment comprès entre aquests dos punts pertany en la seva totalitat al conjunt.)



Convex



No convex

Definició. Combinació convexa.

Siguin $x^1, \dots, x^k \in \mathbb{R}^n$ i $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tq $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \geq 0$, $i = 1, \dots, k$

Llavors, $x = \sum_{i=1}^{k} \lambda_i x^i$ és la combinació convexa de $x^1, \dots, x^k$

És a dir, $x = \lambda_1 x^1 + \dots + \lambda_k x^k$ (tots els x que es poden escriure com a c.l. on $\sum_{i=1}^{k} \lambda_i = 1$ i tots els coeficients $\lambda_i \geq 0$ $i = 1, \dots, k$)

Definició. Embolcall convex. (convex hull)

És el conjunt de totes les combinacions convexes de $x^1, \dots, x^k$ (inclou totes les comb. convexes)

$CH(x^1, \dots, x^k) = \{ x \in \mathbb{R}^n \mid x = \sum_{i=1}^{k} \lambda_i x_i,\ \lambda_i \geq 0\ \sum_{i=1}^{k} \lambda_i = 1 \}$

Propietats (dels conjunts convexos)

a) La intersecció de conjunts convexos és convexa.

b) Tot poliedre és un conjunt convex. ($\Rightarrow$ la regió factible dels problemes PL és un conjunt convex)

c) La combinació convexa d'un nombre finit d'elements d'un conjunt convex pertany al conjunt convex.

d) L'embolcall convex d'un conjunt finit de vectors és un conjunt convex

⊛ La demostració d'aquestes propietats la farem a PROBLEMES

Definició. Punt extrem

Sigui el poliedre P. Un vector $x \in P$ és un punt extrem de P si no existeix cap parell de vectors $y, z \in P$ diferents de x ni cap escalar $\lambda \in [0,1]$ tals que $x = \lambda y + (1-\lambda) z$. ⚠ No tots els poliedres tenen punts extrems

Definició. Línia.

Direm que el poliedre $P \subset \mathbb{R}^n$ conté una línia si existeix el vector $x \in P$ i el vector no nul $d \in \mathbb{R}^n$ tals que $x + \lambda d \in P$ $\forall \lambda \in \mathbb{R}$.

Teorema① Sigui el políedre no buit $P = \{x \in \mathbb{R}^n \mid a'_i x \geq b_i,\ i = 1, \dots, m\}$. Llavors P té com a mínim un punt extrem $\Leftrightarrow$ P no conté cap línia.

Corol·laris.

i) Tot políedre no buit i fitat (polítop) té, com a mínim, un punt extrem.

ii) Tot políedre en forma estàndard no buit té, com a mínim, un punt extrem (ja que $P_e \subset \{x \in \mathbb{R}^n \mid x \geq 0\}$).

Teorema② Teorema Fonamental de la Programació Lineal.

Sigui (PL) $\min_{x \in \mathbb{R}^n} z = \{c'x \mid x \in P\}$, P políedre. Suposem que P conté algun punt extrem i que existeix una solució òptima. Llavors existeix una solució òptima que és un punt extrem de P.

Demostració

1) Demostrem que el conjunt solució $P^* \neq \emptyset$ (conjunt de solucions òptimes) és un políedre que té punts extrems.

- $P^* = \{x \in \mathbb{R}^n \mid \underbrace{Ax \geq b}, \underbrace{c'x = z^*}\}$ $\Rightarrow$ OK políedre

$Ax \geq b$: per la definició del políedre P = regió factible.

$c'x = z^*$: valor òptim de $c'x = z$

$c'x = z^* \equiv c'x \geq z^*$ ''' $-c'x \geq -z^*$

- $P^* \subseteq P$ (regió factible)

hem de veure que no té cap línia.

$\left.\begin{array}{l} P \text{ no conté línees (per l'enunciat)} \\ P^* \subseteq P \end{array}\right\} \Rightarrow P^*$ no té línees $\Rightarrow P^*$ té punts extrems OK

2) Demostrem que si $x^* \in P^*$ és punt extrem de P^*, aleshores x^* és punt extrem de P.

- Reducció a l'absurd. Considerem $x^* \in P^* \mid x^*$ és punt extrem de P^* però $x^* \in P \mid x^*$ NO és punt extrem de P.

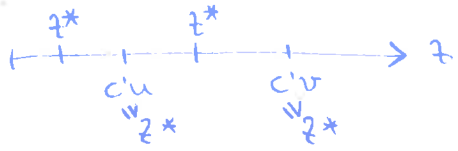
Sigui $u, v \in P \setminus \underbrace{P^*}$, $x^* \neq u, v$ tq $x^* = \lambda u + (1-\lambda)v$, $\lambda \in (0,1)$

redundant, ja que si x^* és pt extrem de P^* i es pot escriure com a c.l. convexe de u i v vol dir que $u, v \notin P^*$

$c'x^* = \lambda \underbrace{c'u}_{\geq z} + (1-\lambda)\underbrace{c'v}_{\geq z} = z^* \Rightarrow c'u = c'v = z^* \Rightarrow$

$\Rightarrow u, v \in P^* \Rightarrow x^*$ NO és punt extrem de P^* CONTRADICCIÓ

Gràficament:

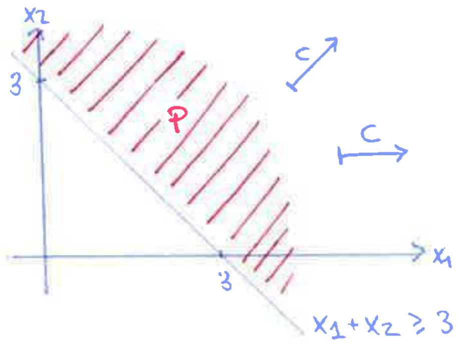


La c.l. de $c'u$ i $c'v$ és z^*

Per tant $c'u$ i $c'v$ han de ser z^*.

⊛ Aplicació del teorema

a)



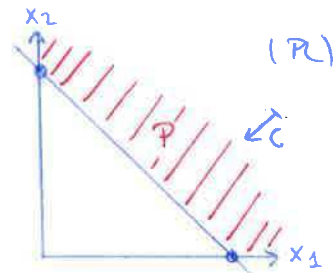
En aquest cas no podem aplicar el teorema perquè no hi ha punts extrems.
↳ Això NO vol dir que no hi hagi solució òptima.

- Sigui $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ el vector gradient de la funció objectiu, aleshores no hi ha solució òptima.
- Sigui $c' = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ el vector gradient d'una altra funció objectiu, llavors tenim infinites solucions òptimes. Són tots els punts de la recta $y=-x+3$

b) Problema infactible. És aquell que viola les hipòtesis de l'enunciat.

$$(PL) \quad \min_{x \in \mathbb{R}^2} \{ c'x \mid x_1 + x_2 \leq 1,\ x_1 + x_2 \geq 3 \}$$

c)



$$(PL) \quad \min_{x \in \mathbb{R}^2} \{ -x_1 - x_2 \mid x_1 + x_2 \geq 3,\ x \geq 0 \}$$

En aquest cas tenim punts extrems però no hi ha òptim

7.- Solucions bàsiques factibles

Comentari. La caracterització de les solucions òptimes de problemes (PL) com a punts extrems NO permet el seu tractament computacional. →
→ solucions bàsiques factibles.

Definició. Solució bàsica (SB):

Sigui $(PL)e \quad \min_{x \in \mathbb{R}^n} \{ c'x \mid Ax = b,\ x \geq 0 \}$, A matriu $m \times n$ de rang complet.

El vector $x \in \mathbb{R}^n$ és una solució bàsica de $(P)e \iff Ax = b$ i existeix un conjunt d'índexs de variables $\beta = \{B(1), \dots, B(m)\}$ tals que:

i) la matriu bàsica $B \overset{\text{def.}}{=} [A_{B(1)}, \dots, A_{B(m)}]$ és no singular

ii) Si $i \notin \beta$ llavors $x_i = 0$.

Exemple.

$$(PL)\left\{\begin{array}{l}\min\ c'x\\ \text{s.a.:}\quad x_1+x_2\le 2\\ \qquad\quad x_1\le 1\\ \qquad\quad x\ge 0\end{array}\right. \xrightarrow{\text{canvi a forma estàndard}} (PL)e\left\{\begin{array}{l}\min\ c'x\\ \text{s.a.:}\quad x_1+x_2+x_3=2\\ \qquad\quad x_1+x_4=1\\ \qquad\quad x\ge 0\end{array}\right.$$

Per tant tenim la matriu $A=\begin{pmatrix}1&1&1&0\\1&0&0&1\end{pmatrix}$ amb $\underbrace{x_1\ x_2}_{\text{var. bàsiques}}$ $\underbrace{x_3\ x_4}_{\text{var. afegides}}$

1.- $\beta=\{1,2\}$, $x_B=\begin{pmatrix}x_1\\x_2\end{pmatrix}$ variables bàsiques $\rightarrow$ $B=\begin{pmatrix}1&1\\1&0\end{pmatrix}$ matriu NO singular

$$B=\begin{pmatrix}1&1\\1&0\end{pmatrix}=(A_1,A_2)=\begin{pmatrix}1x_1&1x_2\\1x_1&0x_2\end{pmatrix}$$

$N=\{3,4\}$, $x_N=\begin{pmatrix}x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ variables no bàsiques $=0$ } si $i\notin B$ llavors $x_i=0$

Per tant, $Ax=b \Rightarrow Ax=\underset{\begin{pmatrix}1&1\\1&0\end{pmatrix}}{B}x_B+\cancel{\underset{\begin{pmatrix}1&0\\0&1\end{pmatrix}}{A_N}x_N}=b \Rightarrow Bx_B=b \Rightarrow \boxed{x_B=B^{-1}\cdot b}$

0

Fem els càlculs: $\beta=\{1,2\}\rightarrow Bx_B=b\Rightarrow \begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2\\1\end{pmatrix}\Rightarrow$

$\Rightarrow \begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0&1\\1&-1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$ Per tant, solució $\boxed{x_B=\begin{pmatrix}1\\1\end{pmatrix}\quad x_N=\begin{pmatrix}0\\0\end{pmatrix}}$

Definició. Solució bàsica factible (SBF).

És una solució bàsica tal que $x\ge 0$.

⊛ Número de possibles solucions bàsiques factibles ($=$ nº de columnes que es poden permutar de A en A) $=\binom{m}{n}$

Exemple. Seguint amb l'exemple d'abans, a part de la solució bàsica que hem trobat podem agafar altres variables com a bàsiques i trobar noves solucions:

2.- $\beta=\{1,4\}\rightarrow B=\begin{pmatrix}1&0\\1&1\end{pmatrix}\Rightarrow \underset{x_B}{\begin{pmatrix}x_1\\x_4\end{pmatrix}}=\underset{B^{-1}}{\begin{pmatrix}1&0\\-1&1\end{pmatrix}}\cdot\underset{b}{\begin{pmatrix}2\\1\end{pmatrix}}=\begin{pmatrix}2\\-1\end{pmatrix}$; $x_N=\begin{pmatrix}x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$

tenim $x_4=-1<0\Rightarrow$ NO és una solució bàsica factible

3.- $\beta=\{2,3\}\rightarrow B=\begin{pmatrix}1&1\\0&0\end{pmatrix}\Rightarrow \begin{pmatrix}x_2\\x_3\end{pmatrix}=B^{-1}\cdot b$

↳ B té rang 1, és una matriu SINGULAR ↳ NO és invertible.

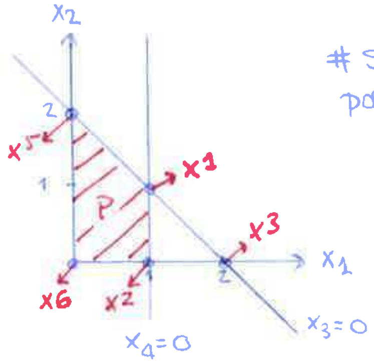
Per tant NO és una solució bàsica.

Definició. Solució bàsica degenerada (SBD).

És una solució bàsica tal que $\exists\, i \in \beta \mid x_i = 0$

Exemple. Ens tornem a fixar en el problema anterior:

$$(PL)_e \begin{cases} \min\ c'x \\ \text{s.a.:} \quad x_1 + x_2 + x_3 = 2 \\ \qquad\quad x_1 + x_4 = 1 \\ \qquad\quad x \geq 0 \end{cases}$$



$\#\,SBF = \binom{4}{2} = 6$ possibles

Analitzem les possibles solucions bàsiques factibles:

$x^1: \beta = \{1,2\} \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad x_N = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

$x^2: \beta = \{1,3\} \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad x_N = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

$x^3: \beta = \{1,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \rightarrow$ infactible

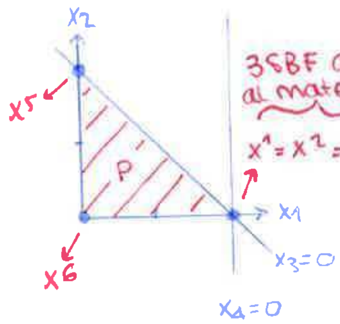
$x^4: \beta = \{2,3\} \rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow B$ matriu singular. $\Rightarrow$ Només 5 S.B.

$x^5: \beta = \{2,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad x_N = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

$x^6: \beta = \{3,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad x_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

Exemple.

$$(PL) \begin{cases} \min\ c'x \\ \text{s.a.} \quad x_1 + x_2 \leq 2 \\ \qquad\ \ x_1 \leq 2 \\ \qquad\ \ x \geq 0 \end{cases} \rightarrow (PL)_e \begin{cases} \min\ c'x \\ \text{s.a.:} \quad x_1 + x_2 + x_3 = 2 \\ \qquad\quad x_1 + x_4 = 2 \\ \qquad\quad x \geq 0 \end{cases}$$



3 SBF assoc. al mateix pt extrem

$x^1 = x^2 = x^3$

$x^1: \beta = \{1,2\} \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow$

$\rightarrow x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad x_N = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

* Tenim 3 variables $= 0 \Rightarrow$ S.B. Degenerada.

$x^2: \beta = \{1,3\} \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow$

$\#$ possibles SBF $= \binom{4}{2} = 6$

$\rightarrow x_B = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad x_N = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBD.

$x^3: \beta = \{1,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_1 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad x_N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBD

$x^4: \beta = \{2,3\} \rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow B$ matriu singular $\Rightarrow$ Només 5 S.B.

$x^5: \beta = \{2,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad x_N = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

$x^6: \beta = \{3,4\} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rightarrow x_B = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad x_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ SBF

Teorema (2.1). Sigui P_e un poliedre estàndard no buit amb $A \in \mathbb{R}^{m \times n}$ i files $a'_1, \dots, a'_m$. Suposem que $rg(A) = k < m$ i que les files $a'_{i_1}, \dots, a'_{i_k}$ són l.i. Considerem el poliedre $Q_e = \{x \in \mathbb{R}^n \mid a'_{i_1} x = b_{i_1}, \dots, a'_{i_k} x = b_{i_k}, x \geq 0\}$. Aleshores $Q_e = P_e$.

↳ És a dir, sigui $\underset{A \in \mathbb{R}^{m \times n}}{A} = \left(\begin{array}{c} \\ \cdots\cdots\cdots \\ \\ \end{array} \right) \begin{array}{l} a_{i_1} \\ a_{i_k} \\ a_{i_m} \end{array}$ } files l.i. $\Rightarrow$ prenem $\underset{\tilde{A} \in \mathbb{R}^{k \times n}}{\tilde{A}} = \left(\begin{array}{c} \\ \\ \end{array} \right) \begin{array}{l} a_{i_1} \\ a_{i_k} \end{array}$

$rg(\tilde{A}) = k \rightarrow$ totes les files l.i.

Aleshores, $P_e = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\} \equiv Q_e = \{x \in \mathbb{R}^n \mid \tilde{A} x = \tilde{b},\ x \geq 0\}$

$A \in \mathbb{R}^{m \times n}$, $rg(A) = k$ $\qquad$ $\tilde{A} \in \mathbb{R}^{k \times n}$, $rg(\tilde{A}) = k$

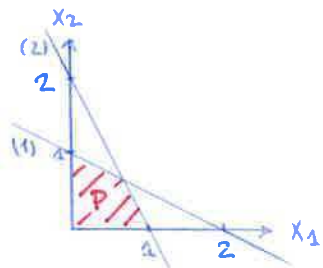
El teorema ens diu que si tenim una matriu A de rang NO complet i amb solucions factibles, si agafem les files l.i. tindrem una matriu de rang complet i el poliedre serà el mateix.

Demostració → Exercici 11 de problemes.

La idea és veure que si una fila és c.l. de les altres, qualsevol punt que satisfaci aquella constricció també satisfarà la c.l. de les altres.

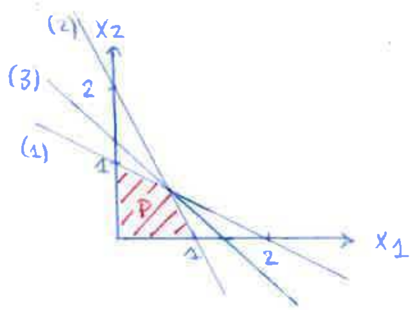
Exemple.

$$(PL)_e \begin{cases} \min\ c'x \\ \text{s.a.:}\quad \frac{1}{2}x_1 + x_2 = 1 \quad (1) \\ \qquad x_1 + \frac{1}{2}x_2 = 1 \quad (2) \\ \qquad x \geq 0. \end{cases}$$



afegim una tercera restricció que sigui c.l. de les altres dues

$$(PL)_e \begin{cases} \min\ c'x \\ \text{s.a.:}\quad \frac{1}{2}x_1 + x_2 = 1 \\ \qquad x_1 + \frac{1}{2}x_2 = 1 \\ \qquad \frac{3}{2}x_1 + \frac{3}{2}x_2 = 2 \\ \qquad x \geq 0 \end{cases}$$



La tercera restricció és redundant.

⊛ Així doncs, si A no és de rang complet i P_e no és buit, les files l.d. de A tenen associades constriccions redundants i es poden eliminar de la formulació del problema sense que la regió factible canviï.

Teorema ③. Sigui P un poliedre no buit en forma estàndard de rang complet, i sigui $x^* \in P$. Aleshores x^* és un punt extrem $\Leftrightarrow$ x^* és una SBF.

Demostració

1.- x punt extrem $\Rightarrow$ x SBF

a.- $x = [x_1, \dots, x_r, 0, \dots, 0]$ punt extrem de P_e amb $x_1, \dots, x_r > 0$

$Ax = b \Rightarrow \sum_{i=1}^{r} A_i x_i = b \quad (1)$

b.- Volem veure que $\sum_{i=1}^{r} A_i x_i = b$ són l.i.

Reducció a l'absurd:

$A_i \quad i = 1, \dots, r$ són l.d. ← suposem.

Sigui x punt extrem i $\exists \alpha_i \neq 0 \quad \sum_{i=1}^{r} A_i \alpha_i = 0 \quad (2)$

Amb (1) i (2) i prenent $\varphi > 0$,

$$\sum_{i=1}^{r} A_i (x_i + \varphi \alpha_i) = b \; , \quad \sum_{i=1}^{r} A_i (x_i - \varphi \alpha_i) = b \quad (3)$$

$Ax^1 = b$ $\qquad$ ja que $\varphi \cdot \sum_{i=1}^{r} A_i \alpha_i = 0$ $\qquad$ $Ax^2 = b$

Tenint en compte que $x_i > 0, \; \alpha_i \neq 0 \quad i = 1, \dots, r \overset{(3)}{\Rightarrow}$

$\overset{(3)}{\Rightarrow} \; \exists \tilde{\varphi} > 0 \mid \forall \varphi \in [0, \tilde{\varphi}] \mid x_i + \varphi \alpha_i > 0, \; x_i - \varphi \alpha_i > 0 \quad (4)$
$\qquad i = 1, \dots, r$

Definim $x^1 = [x_1 + \varphi \alpha_1, \dots, x_r + \varphi \alpha_r, 0, \dots, 0]'$

vectors $\quad x^2 = [x_1 - \varphi \alpha_1, \dots, x_r - \varphi \alpha_r, 0, \dots, 0]'$

$x^1, x^2 \in P_e$ (3) $\quad ; \quad x^1, x^2 \geq 0$ (4)

$\frac{1}{2} x^1 + \frac{1}{2} x^2 = x$, però x punt extrem $\Rightarrow$ CONTRADICCIÓ!! $\Rightarrow$

$\Rightarrow A_i, \; i = 1, \dots, r$ són l.i.

$\Rightarrow$ x SBF tal que podem distingir dos casos

(següent full).

Cas A

x SBF | r = m   (tenim tantes variables NO nules com files té la matriu A. La matriu bàsica està formada per $A_1, \ldots, A_r = m$).

↓

el vector x és el que hem definit com a $\underbrace{\text{sol. bàsica}}$ $\underbrace{\text{factible}}$

perquè aquest punt està format per m variables associades a una matriu B.

punt extrem

Cas B

x SBF | r < m   (puc trobar una matriu B NO singular).

rg A = m $B = [A_1, \ldots, \underbrace{A_r, \ldots, A_m}_{\text{les afegim}}]$

↓

solució bàsica factible degenerada →

$$m\left[\left(\begin{array}{c|c|c} 1 \ldots & r \ldots & m \end{array}\quad\right)\right] \quad \text{rg } A = m.$$

B

A

puc trobar columnes que em permetin completar aquesta base no singular.

2.- x SBF ⇒ x punt extrem

a.- $x \in Pe$, SBF $x = [x_1, \ldots, x_s, 0, \ldots, 0]$ $x_j > 0$ $j = 1, \ldots, s$ $s \leq m$ (1)

b.- $\underbrace{\sum_{i=1}^{s} A_i x_i = b}_{Ax = b}$ A_i $i = 1, \ldots, s$ són l.i.

c.- Demostrarem per reducció a l'absurd que x és punt extrem.

Considero x NO punt extrem. ⇒ $x = \lambda x^1 + (1-\lambda) x^2$; $\lambda \in (0,1)$ $x^1, x^2 \in Pe$ $x^1, x^2 \neq x$ (2)

$x^1, x^2 \geq 0$ ⇒ $x_i^1 = x_i^2 = 0$ $i = s+1, \ldots, n$ (3)   les últimes n-s components són zero.

$x, x^1, x^2 \in Pe \overset{(3)}{\Longrightarrow} \sum_{i=1}^{s} A_i x_i = \sum_{i=1}^{s} A_i x_i^1 = \sum_{i=1}^{s} A_i x_i^2 = b$ (4)

$\left.\begin{array}{l} A_i \quad i = 1, \ldots, s \quad \text{l.i.} \\ \quad (4) \end{array}\right\} \Rightarrow \left.\begin{array}{l} x_i = x_i^1 = x_i^2 \quad i = 1, \ldots, s \\ \quad (3) \end{array}\right\} \Rightarrow x = x^1 = x^2 \Rightarrow$

⇒ x és punt extrem → CONTRADICCIÓ!!

per tant x SBF ⇒ x punt extrem OK

Solucions Bàsiques: càlcul.

- **Càlcul d'una SB**: considerem el següent políedre:

$$P = \left\{ x \in \mathbb{R}^2 \,\middle|\, \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} x \le \begin{bmatrix} 6 \\ 0 \end{bmatrix}, x \ge 0 \right\}$$

1. Es sel·leccionen m variables d'indexos $\mathcal{B} = \{B(1), \dots, B(m)\}$ amb columnes de A linealment independents (variables bàsiques):

$$x_B = [x_{B(1)}, x_{B(2)}, \dots, x_{B(m)}]' \; , B \stackrel{\text{def}}{=} [A_{B(1)}, A_{B(2)} \dots, A_{B(m)}]$$

2. Es fixen les $n - m$ variables no bàsiques a zero:

$$\mathcal{N} \stackrel{\text{def}}{=} \{i \mid i \notin \mathcal{B}\} \equiv \{N(1), \dots, N(n-m)\}$$

$$x_N = [0] \quad , A_N \stackrel{\text{def}}{=} [A_{N(1)}, A_{N(2)}, \dots, A_{N(n-m)}] \text{ (matriu no-bàsica)}$$

3. Es calcula el valor de les variables bàsiques:

$$Ax = [B \quad A_N] \begin{bmatrix} x_B \\ x_N \end{bmatrix} = Bx_B + A_N \overbrace{x_N}^{=0} = Bx_B = b, \boxed{x_B = B^{-1} b}$$

Conjunt de les SB d'un problema PL

- **Problema de planificació de la producció en forma estàndard:**

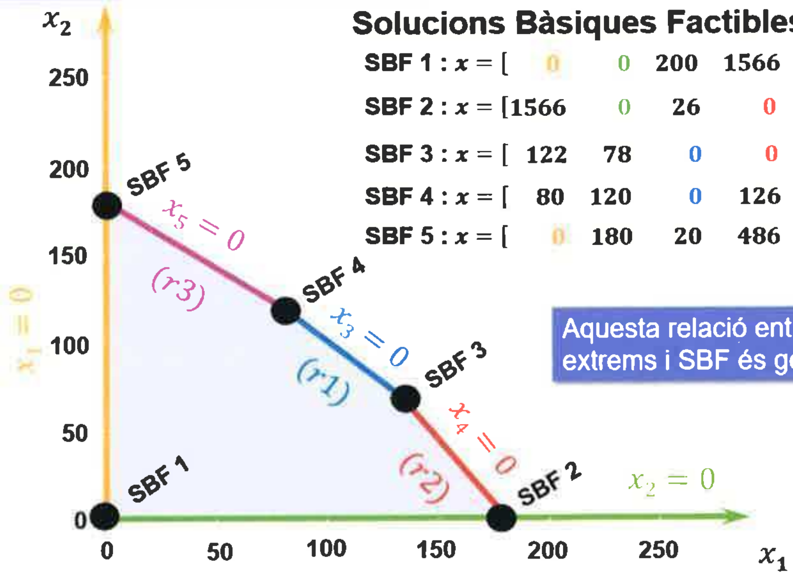
$$(PL) \begin{cases} \min & z = -350x_1 \quad -300x_2 \\ \text{s.a.:} & x_1 \quad +x_2 \quad +x_3 \quad = 200 \quad (r1) \\ & 9x_1 \quad +6x_2 \quad +x_4 \quad = 1566 \quad (r2) \\ & 12x_1 \quad +16x_2 \quad +x_5 \quad = 2880 \quad (r3) \\ & x_1 \ge 0, \quad x_2 \ge 0, \quad x_3 \ge 0, \quad x_4 \ge 0 \quad x_5 \ge 0 \end{cases}$$

- **Nombre total de SB** $\le \binom{n}{m} = \binom{5}{3} = 5!/3!\,2! = 10$

SB	x_B	x_N	x	z
1	x_3, x_4, x_5	x_1, x_2	$x = [0 \quad 0 \quad 200 \quad 1566 \quad 2880]'$	0
2	x_1, x_3, x_5	x_2, x_4	$x = [174 \quad 0 \quad 26 \quad 0 \quad 792]'$	-60900
3	x_1, x_2, x_5	x_3, x_4	$x = [122 \quad 78 \quad 0 \quad 0 \quad 168]' = x^*$	-66100
4	x_1, x_2, x_4	x_3, x_5	$x = [80 \quad 120 \quad 0 \quad 126 \quad 0]'$	-64000
5	x_2, x_3, x_4	x_1, x_5	$x = [0 \quad 180 \quad 20 \quad 486 \quad 0]'$	-54000
6	x_1, x_2, x_3	x_4, x_5	$x = [108 \quad 99 \quad -7 \quad 0 \quad 0]'$	-67500
7	x_1, x_3, x_4	x_2, x_5	$x = [240 \quad 0 \quad -40 \quad -594 \quad 0]'$	-84000
8	x_1, x_4, x_5	x_2, x_3	$x = [200 \quad 0 \quad 0 \quad -234 \quad 480]'$	-70000
9	x_2, x_4, x_5	x_1, x_3	$x = [0 \quad 200 \quad 0 \quad 366 \quad -320]'$	-60000
10	x_2, x_3, x_5	x_1, x_4	$x = [0 \quad 261 \quad -61 \quad 0 \quad -1296]'$	-78300

Solucions bàsiques factibles — solució bàsica factible òptima — solucions bàsiques infactibles

Solucions Bàsiques Factibles i punts extrems



Solucions Bàsiques Factibles :

SBF 1 : $x = [\ 0 \quad 0 \quad 200 \quad 1566 \quad 2880]'$

SBF 2 : $x = [1566 \quad 0 \quad 26 \quad 0 \quad 792]'$

SBF 3 : $x = [\ 122 \quad 78 \quad 0 \quad 0 \quad 168]'$

SBF 4 : $x = [\ 80 \quad 120 \quad 0 \quad 126 \quad 0]'$

SBF 5 : $x = [\ 0 \quad 180 \quad 20 \quad 486 \quad 0]'$

Aquesta relació entre pts. extrems i SBF és general.

TEMA 2 : Algorisme del símplex

1.- Introducció

Com justifiquem l'algorisme del símplex?

Recordem: <u>Teorema 2</u> → optimalitat pts. extrems
<u>Teorema 3</u> → equivalència pts. extrems - SBF } ⇒

⇒ **Corol·lari** (optimalitat SBF)

Sigui (PL) $\min\{c'x \mid x \in P\}$, P poliedre. Suposem que P conté algun punt extrem i que existeix una solució òptima. Aleshores existeix una solució òptima que és una SBF de P. → **Això ens dóna una via per a resoldre problemes PL.**

⊛ Idea de l'algorisme del símplex:

Passem d'una SBF x a un altre y tal que $c'y < c'x$, fins a trobar l'òptima.

⊛ Questions a resoldre:

1.- Com trobem una SBF? **Fase I del simplex**

2.- Com canviem d'una SBF a un altre millor? **Direccions bàsiques factibles de descens**

3.- Com identifiquem la SBF òptima? **Condicions d'optimalitat**

4.- I si el problema no té solució? **Problemes il·limitats**

2.- Direccions bàsiques factibles (DBF)

Definició. SBF adjacents

Dues SBF són adjacents si es distingeixen només per una variable bàsica.

Definició. Direcció factible

La direcció factible és el vector $d \in \mathbb{R}^n$ factible sobre $x \in P_e$ si $\exists\, \theta \in \mathbb{R}^+$ tal que $x + \theta d \in P_e$

Definició. Direcció bàsica factible (DBF)

La direcció bàsica factible sobre la SBF $x \in P_e$ associada a $q \in N$ és la direcció $d = \begin{pmatrix} d_B \\ d_N \end{pmatrix} \in \mathbb{R}^n$ tal que:

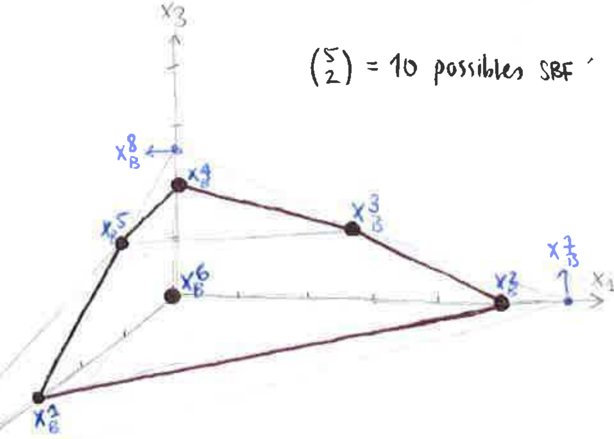
i) $d_{N(i)} = \begin{cases} 1 & \text{si } N(i) = q \\ 0 & \text{si } N(i) \neq q \end{cases}$ $\quad i = 1, \dots, n-m$

ii) d_B tal que $A(x + \theta d) = b$ per a algun $\theta > 0 \Rightarrow d_B = -B^{-1} \cdot A_q$.

Exemple.

$\binom{5}{2} = 10$ possibles SBF

$$(PL)\begin{cases} \min \quad z = 5x_1 + 3x_2 + 9x_3 \\ \text{s.a.:} \\ \quad x_1 + x_2 + 3x_3 + x_4 = 6 \\ \quad 5x_1 + 3x_2 + 6x_3 + x_5 = 15 \\ \quad x \geq 0 \end{cases}$$



Sigui SBF x^1 i les seves SBF adjacents:

x^1: $B = \{1,4\}$, $x_B^1 = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = B^{-1} \cdot b = \begin{bmatrix} 1 & 1 \\ 5 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 15 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $\quad z^1 = c'_B x_B = 15$

x^2: $B = \{2,4\}$, $x_B^2 = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = B^{-1} \cdot b = \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 15 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ $\quad z^2 = c'_B x_B = 15$

x^5: $B = \{1,3\}$, $x_B^5 = \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = B^{-1} \cdot b = \begin{bmatrix} 1 & 3 \\ 5 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 15 \end{bmatrix} = \begin{bmatrix} 1 \\ 5/3 \end{bmatrix}$ $\quad z^5 = c'_B x_B = 20$

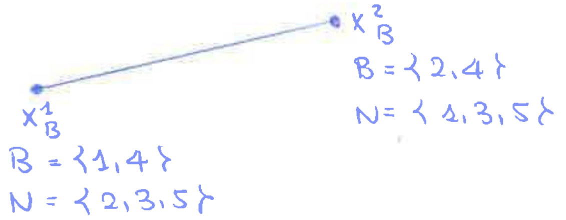
x^6: $B = \{4,5\}$, $x_B^6 = \begin{bmatrix} x_4 \\ x_5 \end{bmatrix} = B^{-1} \cdot b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 15 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix}$ $\quad z^6 = c'_B x_B = 0$

x^9: és una solució bàsica infactible (tenim una variable bàsica < 0)

Ara desenvolupem el procediment per a passar de la SBF x^1 a qualsevol de les seves SBF adjacents.

✱ Sigui P_e, no buit, rang complet, no degenerat. Donades les SBF adjacents x i y, volem trobar la direcció $d \in \mathbb{R}^n$ i l'escalar $\theta^* \in \mathbb{R}^+$ tal que $y = x + \theta^* d$ → Així doncs, he d'imposar que $\begin{cases} A \cdot y = b \\ y \geq 0 \end{cases} \in P_e$

En el nostre exemple tenim:



Per tant, per passar de x_B^1 a x_B^2 es compleix que $q = 2$, $d_N = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{matrix} \leftarrow d_2 \\ \leftarrow d_3 \\ \leftarrow d_5 \end{matrix}$

Per una altra banda tenim $d_B = \begin{bmatrix} d_1 \\ d_4 \end{bmatrix} = -B^{-1} \cdot A_{\underset{2}{q}} = -\begin{pmatrix} 0 & 1/5 \\ 1 & -1/5 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -3/5 \\ -2/5 \end{pmatrix}$

✱ La DBF d és factible sobre la SBF $x \in P_e$ si $\exists \theta \in \mathbb{R}^+ \mid y = x + \theta d \in P_e$

θ^* és la longitud de pas quan passem d'una SBF a una altra.

$\theta^* = \max\{\theta > 0 \mid y = x + \theta d \in P_e\}$ → d'aquesta manera arribem a un altre punt extrem.

↓ longitud de pas màxima

Ara farem el càlcul de θ^* (longitud de pas màxima).

$y \in Pe = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$ tq. 1) $Ay = A(x + \theta d) = b$, cert $\forall \theta$

2) $y = x + \theta d \geq 0$, depèn de θ:

- $y_{N_i} = \begin{cases} 0 & \text{si } i \neq q \\ \theta & \text{si } i = q \end{cases} \geq 0 \quad \forall \theta > 0.$

- $y_B = \underbrace{x_B}_{>0} + \underbrace{\theta d_B}_{?} \geq 0$, llavors $y_B \geq 0 \iff \theta \leq \theta^*$ amb:

$$\theta^* = \max\{\theta > 0 \text{ tq. } x + \theta d \geq 0\} = \min_{\{i=1,\dots,m \mid d_{B(i)} < 0\}} \{-x_{B(i)}/d_{B(i)}\}$$

En el nostre exemple:

$$y_B = x_b + \theta d_B \Rightarrow y_B = \begin{bmatrix} 3 \\ 3 \end{bmatrix} + \theta \begin{bmatrix} -3/5 \\ -2/5 \end{bmatrix} \geq 0 \Rightarrow \theta \leq \theta^*,$$

on $\theta^* = \min_{\{d_{B(1)}, d_{B(4)}\}} \left\{\frac{-3}{-3/5}, \frac{-3}{-2/5}\right\} = \min\{5, \frac{15}{2}\} = 5.$

Un cop tenim la d i la θ^*, ja podem calcular la nova SBF:

$$y = x + \theta^* d = \begin{pmatrix} x_B \\ x_N \end{pmatrix} + \theta^* \begin{pmatrix} d_B \\ d_N \end{pmatrix}$$

$$x^2 = x^1 + \theta^* d$$

$$x^2 = \begin{array}{c} x_1 \\ x_4 \\ x_2 \\ x_3 \\ x_5 \end{array} \begin{bmatrix} 3 \\ 3 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} -3/5 \\ -2/5 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} \Big\}$$ les variables bàsiques de x^2 són x_4 i x_2.

SBF adjacents.

El canvi de base efectuat és: $x^1 = \{1, 4\} \rightarrow x^2 = \{2, 4\}$

Propietats.

i) Si Pe és <u>no</u> degenerat ($x_{B(i)} > 0\ \forall i \in \beta$) d és factible:

a) Si $d_B \not\geq 0 \Rightarrow \theta^* > 0 \Rightarrow d$ factible

b) Si $d_B \geq 0 \Rightarrow y_{B(i)} = x_{B(i)} + \theta d_{B(i)} \geq 0\ \forall \theta > 0$

llavors θ no està definida, doncs $\forall \theta > 0$ $x + \theta d \in Pe$ (d és un raig extrem)

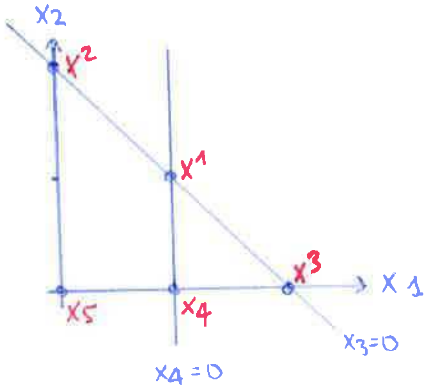
ii) Si Pe és degenerat ($\exists i \in \beta \mid x_{B(i)} = 0$) d pot no ser factible:

Si $\exists i \in \beta \mid x_{B(i)} = 0$ i $d_{B(i)} < 0$

llavors $\min_{\{i=1,\dots,m \mid d_{B(i)} < 0\}} \{-x_{B(i)}/d_{B(i)}\} = 0 \Rightarrow \nexists\, \theta > 0 \mid y = x + \theta d \in Pe \Rightarrow d$ infactible.

Exemple.

$$(PL)\begin{cases} \min\ c'x \\ \text{s.a.:}\quad x_1+x_2\le 2 \\ \qquad\quad x\le 1 \\ \qquad\quad x\ge 0\end{cases} \Rightarrow (PL)_e \begin{cases}\min\ c'x \\ \text{s.a.:}\quad x_1+x_2+x_3=2 \\ \qquad\quad x_1+x_4=1 \\ \qquad\quad x\ge 0\end{cases}$$



x^1: $\beta=\{1,2\}$ $\quad B=\begin{pmatrix}1&1\\1&0\end{pmatrix}$ $\quad x_B^1=\begin{pmatrix}1\\1\end{pmatrix}$

x^2: $\beta=\{2,4\}$ $\quad B=\begin{pmatrix}1&0\\0&1\end{pmatrix}$ $\quad x_B^2=\begin{pmatrix}2\\1\end{pmatrix}$

x^3: $\beta=\{1,4\}$ $\quad B=\begin{pmatrix}1&0\\1&1\end{pmatrix}$ $\quad x_B^3=\begin{pmatrix}-1\\-1\end{pmatrix}\rightarrow$ solució NO bàsica

x^4: $\beta=\{1,3\}$ $\quad B=\begin{pmatrix}1&1\\1&0\end{pmatrix}$ $\quad x_B^4=\begin{pmatrix}1\\1\end{pmatrix}$

x^5: $\beta=\{3,4\}$ $\quad B=\begin{pmatrix}1&0\\0&1\end{pmatrix}$ $\quad x_B^5=\begin{pmatrix}2\\1\end{pmatrix}$

(x^6: $\beta=\{2,3\}$ $\quad B=\begin{pmatrix}1&1\\0&0\end{pmatrix}\rightarrow$ matriu singular.)

INFACTIBLE.

Ara volem passar de la SBF x^1 a x^2. $\quad x_B^1=\{1,2\}\quad x_B^2=\{2,4\}$

La component associada a la variable 3 és zero perquè x_3 és no bàsica en les dues SBF. $\Rightarrow d_3=0$.

En canvi, la component associada a d_4 ha de ser >0, ja que la variable passa de no bàsica a bàsica $\Rightarrow d_4=+1$.

$$Ax^2=b,\quad \underbrace{Ax^1}_{b}+\theta Ad=b \Rightarrow \underbrace{\theta}_{>0}Ad=0 \Rightarrow Ad=0\ ;\quad d=\begin{bmatrix}d_B\\d_N\end{bmatrix}$$

Tenim que $Ad=0$. Considerem que la matriu A està ordenada tal que $A=[\,B\ \vdots\ A_n\,]$ $\quad$ B: m columnes que corresponen a les VB. $\quad A_n$: $n-m$ columnes que corresponen a les VNB

$$Ad = B\cdot d_B\ ,\ A_n\cdot \underset{\sim}{d_N}=0$$

$$d_N=\begin{cases}0 & \text{si } N(i)\neq q\\ 1 & \text{si } N(i)=q\end{cases}$$

En el nostre cas $q=4$

variable que passa de no bàsica a bàsica

$$B\cdot d_B+\overbrace{A_{N(i)}}^{A_q}=0 \Rightarrow B\,d_B=-A_q \Rightarrow d_B=-B^{-1}\cdot A_q$$

En el nostre exemple: columna de la matriu de restriccions associada a x_4 perquè $q=4$

$$d_B=-\begin{array}{c} \begin{matrix}x_1 & x_2\end{matrix}\\ \begin{bmatrix}1&1\\1&0\end{bmatrix}^{-1}\end{array}\begin{bmatrix}0\\1\end{bmatrix}=-\begin{bmatrix}0&1\\1&-1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-1\\1\end{bmatrix}\begin{matrix}\leftarrow d_1\\ \leftarrow d_2\end{matrix}$$

$$d_N=\begin{bmatrix}0\\1\end{bmatrix}\begin{matrix}\leftarrow d_3\\ \leftarrow d_4\end{matrix}$$

$$d=\begin{bmatrix}-1\\1\\0\\1\end{bmatrix}$$

Així doncs, imposant que x^2 és SBF adjunta a x^1 hem definit l'expressió del vector d que satisfà que $A \cdot x^2 = b \ \forall \theta > 0 \quad (x^2 = x^1 + d\theta)$

Volem veure si $y = x + \theta^* d \geq 0 \quad (\theta > 0)$

$i \in N_x$: $y_i = \underbrace{\cancel{x_i}}_{0} + \theta^* d_i = \begin{cases} \theta^* \cdot 0 = 0 & \text{si } i \neq q \\ \underbrace{\theta^* \cdot 1 > 0}_{\forall \theta^* > 0} & \text{si } i = q \end{cases}$ $\left(d_i = \begin{cases} 0 & i \in N_x, i \neq q \\ 1 & i \in N_x, i = q \end{cases} \right)$

perquè són les variables no bàsiques

↳ per definició.

Hem de veure quines condicions s'han de satisfer perquè d sigui ≥ 0

Tenim $B_x = \{B(1), \dots, B(m)\}$ → En el nostre cas $B_x = \{1, 2\}$

$i \in 1, \dots, m$ tq $y_{B(i)} = \underbrace{x_{B(i)}}_{>0} + \underbrace{\theta^*}_{>0} \underbrace{d_{B(i)}}_{\gtrless 0} \geq 0$ → Distingim casos $\begin{cases} x_{B(i)} + \theta^* d_{B(i)} > 0 & i : d_{B(i)} \geq 0 \\ x_{B(i)} + \theta^* d_{B(i)} \gtrless 0 & i : d_{B(i)} < 0 \end{cases}$

Quina condició s'ha de satisfer per tal de poder assegurar que aquesta actualització sigui ≥ 0? ← volem que això sigui ≥ 0

↳ $\theta^* \leq \frac{-x_{B(i)}}{d_{B(i)}} \quad \forall i$ tq $d_{B(i)} < 0$. → Això s'ha de satisfer per a totes les VB que tinguin la component $d_{B(i)} < 0$

Jo vull el màxim valor de θ^* per tal de no sortir del poliedre

$$\theta^* = \max\{\theta > 0 \mid x + \theta d \geq 0\} = \min_{i=1,\dots,m \mid d_{B(i)} < 0} \left\{ {-x_{B(i)}}/{d_{B(i)}} \right\}$$

En el nostre exemple, $d_B = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow d_1 \\ \leftarrow d_2 \end{matrix} \Rightarrow \theta^* = \min\left\{ \frac{-x_{B(1)}}{d_{B(1)}} \right\} = \min\left\{ \frac{-1}{-1} \right\} = 1$

Així doncs, $y = x + \theta d \Rightarrow x^2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}; \ x_B^2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow x_2 \\ \leftarrow x_4 \end{matrix}$

Factibilitat de d i degeneració de Pe

La nostra d en l'exemple anterior és una DBF, que en els poliedres no degenerats és un cas particular de les direccions factibles.

Anem a veure de què depèn la factibilitat de d quan tenim un poliedre degenerat. → tenim $\overset{VB}{x_B}$ que poden ser zero. ⇒ SBF degenerada

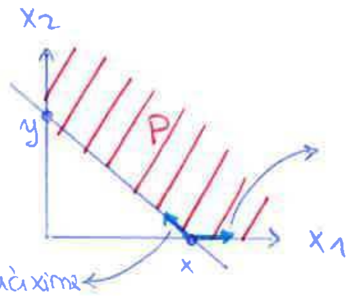
↳ Considerem un poliedre no buit i de rang màxim i imposem que $x_i \geq 0$, $i = 1, \dots, m$ → si es dóna el cas que tenim una variable bàsica que és zero, aleshores $\min\left\{ \frac{-x_{B(i)}}{d_{B(i)}} \right\} = \min\left\{ \frac{0}{d_{B(i)}} \right\} = \boxed{0 = \theta^*}$

d'acord amb la definició de DBF, la passa màxima que podem fer (θ^*) és zero → això no compleix que $\theta > 0$

En canvi, si sabem que P és **no** degenerat $\Rightarrow$ totes les variables bàsiques són > 0 $\Rightarrow$ tot el desenvolupament que hem fet anteriorment és vàlid.

⚠ Pot ser que trobem una d on totes les seves components són positives. Com interpretem això? Ens podem moure tal com vulguem al llarg d'aquella direcció i mai sortirem del poliedre.
(No tenim fita superior per a θ^*)

Exemple



Tenim 2 DBF sobre x

No tenim una θ^* màxima
d és un raig extrem

tenim θ^* màxima definida.

✻ Si q i $B(p)$ representen, respectivament, les variables que entren i surten de la base, llavors $\bar{B} := \{\bar{B}(1), \dots, \bar{B}(m)\}$ amb $\bar{B}(i) := \begin{cases} B(i) & i \neq p \\ q & i = p \end{cases}$

; la nova base és $\bar{B} = [A_{B(1)}, \dots, A_{B(p-1)}, A_q, A_{B(p+1)}, \dots, A_{B(m)}]$

↳ És a dir, la variable que entra substitueix a la que surt, sense reordenar-ho pels índexs

Ara veurem que $\bar{B}$ és SBF.

Teorema. ④ Sigui x SBF de P_e poliedre estàndard de rang complet, no buit i no degenerat, d DBF sobre x i $\theta^* = \max\{\theta > 0 \mid y = x + \theta d \in P_e\}$. Aleshores $y = x + \theta^* d$ és SBF de P_e.

Demostració

Volem veure que y és SBF. $(y = x + \theta^* \cdot d)$ $\bar{B}$ és la base de y.

- Per construcció sabem que $y \in P_e$ i té $n-m$ variables no bàsiques nul·les
$y = [\underbrace{y_B}_{m} \,\vdots\, \underbrace{\overset{0}{\overset{\shortparallel}{y_N}}}_{n-m}] = \bar{B}$ $\rightarrow$ hem de veure que $\bar{B}$ és no singular $\Rightarrow$ y és SBF

- Per reducció a l'absurd. Considerem $\bar{B}$ singular $\rightarrow$ arribarem a contradicció

$\bar{B}$ singular $\Rightarrow \exists \lambda_1, \dots, \lambda_m$ no tots $\neq 0$ tq $\sum_{i=1}^{m} \lambda_i A_{\bar{B}(i)} = 0$ $\Rightarrow$

$\Rightarrow B^{-1} \cdot \sum_{i=1}^{m} \lambda_i A_{\bar{B}(i)} = \sum_{i=1}^{m} \lambda_i B^{-1} \cdot A_{\bar{B}(i)} = 0.$

Veiem que $B^{-1} \cdot A_{\bar{B}(i)}$ són l.i. per arribar a una contradicció:

- Si $i \neq p$; $B^{-1} \cdot A_{\bar{B}(i)} = B^{-1} \cdot A_{B(i)}$ l.i. ok.
- Si $i = p$, $B^{-1} \cdot A_q = -d_B$: per definició $\begin{bmatrix} 1 \\ p \\ m \end{bmatrix}$ p - $\overset{<0}{\overbrace{d_{B(p)}}} > 0 \Rightarrow B^{-1}(A_{\bar{B}(1)}, \dots, A_{\bar{B}(m)}) =$
$= \begin{pmatrix} 1 & & \vdots & & \\ & \ddots & \vdots & & \\ & & >0 & & \\ & & \vdots & \ddots & \\ & & \vdots & & 1 \end{pmatrix}$ columnes L.I
↳ $\bar{B}$ no singular

3.- Condicions d' optimalitat

Siguin $x, y \in P_e$ dues SBF adjacents tq $y = x + \theta^* d$.

$$c'y = \underbrace{c'x}_{B \longrightarrow [c'_B, c'_N]} + \theta^* \cdot c'd = c'x + \theta^*[c'_B \cdot \underbrace{d_B}_{= -B^{-1} \cdot A_q} + \underbrace{c'_N \cdot d_N}_{= c_q}] = c'x + \theta^*(\underbrace{c_q - c'_B B^{-1} A_q}_{r_q}) =$$

$$= c'x + \theta^* r_q$$

r_q ← cost reduït de la variable no bàsica q.

cada VNB té el seu cost reduït

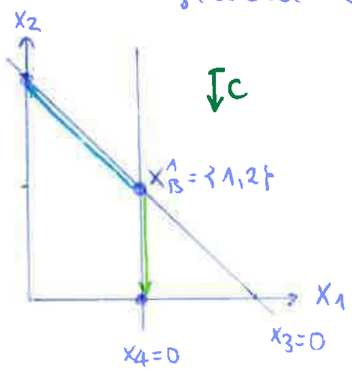
Així doncs tenim $c'y = c'x + \underbrace{\theta^*}_{>0} \underbrace{r_q}_{\text{pot ser } \gtrless 0}$ amb $r_q = c_q - c'_B B^{-1} A_q$ $(q \in N)$

si no degeneració ↲ És un escalar.

Distingim els casos:

- si $r_q < 0 \Rightarrow c'y < c'x$ → millora
- si $r_q = 0 \Rightarrow c'y = c'x$ → no millora
- si $r_q > 0 \Rightarrow c'y > c'x$ → empitjora

Exemple. Imaginem el problema anterior:



→ canvi de variable que ens introdueix x_3

→ canvi de variable que ens introdueix x_4

⊛ Inventem una f. objectiu tq. el seu vector de costos $\underline{c}$ és $c = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}$

Aleshores, tenint $c = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{cases} r_3 = c_3 - [c_1, c_2] B^{-1} A_3 = 0 - [0, -1] \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \\ r_4 = c_4 - [c_1, c_2] B^{-1} A_4 = 0 - [0, -1] \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1 \end{cases}$

(el vector de costos reduïts seria $r = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{matrix} \leftarrow r_3 \\ \leftarrow r_4 \end{matrix}$)

per tant, $r_3 > 0$, $r_4 < 0$ } Això vol dir que el canvi de base introduint x_4 ens millora la nostra f. objectiu, i l'altre canvi de base ens l'empitjora.

Teorema ⓔ Condicions d'optimalitat de SBF.

Sigui P_e poliedre no buit en forma estàndard, x SBF de P_e i sigui el vector de costos reduïts associat a x: $r' = c'_N - c'_B B^{-1} A_N$. Llavors:

a) si $r \geq [0] \Rightarrow x$ és SBF òptima

b) si x és SBF òptima i no degenerada $\Rightarrow r \geq [0]$

Demostració

↳ Idea: considerarem una SBF i un vector factible qualsevol. Demostrarem que quan ens movem en qualsevol direcció factible, si el cost reduït és > 0 sempre empitjorarem la funció.

a) Si $r \geq 0 \Rightarrow x$ és SBF òptima

- Sigui $v \in P_e$ una solució factible qualsevol (no necessàriament SBF)
- Sigui $d = v - x$. $A \cdot d = A(v - x) = Av - Ax = b - b = 0$.

reordenem la matriu

$$A \cdot d = [\underbrace{B}_{\text{NO singular}}, A_N] \begin{bmatrix} d_B \\ d_N \end{bmatrix} = B d_B + A_N d_N = 0 \Rightarrow d_B = -B^{-1} \cdot A_N \cdot d_N$$

- $c'v = c'x + c'd = c'x + [c'_B, c'_N] \begin{bmatrix} d_B \\ d_N \end{bmatrix} =$

$$= c'x + c'_B(-B^{-1} A_N d_N) + c'_N d_N =$$

$$= c'x + d_N (\underbrace{c'_N - c'_B \cdot B^{-1} \cdot A_N}_{r'}) = c'x + \underbrace{r'}_{\geq 0} \cdot \underbrace{d_N}_{?}$$

$d_N = \underbrace{v_N}_{\geq 0 \text{ perquè és sol. factible}} - x_N \geq 0 \Rightarrow r' \cdot d_N \geq 0 \Rightarrow c'v \geq c'x \Rightarrow x$ SBF òptima OK

b) x SBF òptima NO degenerada $\Rightarrow r \geq [0]$

Ho demostrarem per reducció a l'absurd:

- Suposem x SBF òptima NO degenerada i $\exists q \in N \mid r_q < 0$
- Aleshores, sigui d DBF associada a x i q $\rightarrow d_q = e_j \; (N_{(j)} = q)$

  Si x SBF NO degenerada podem assegurar que la DBF d associada a q té $\theta^* > 0$.

  ↗ DBF associada a la part bàsica x i la part no bàsica q
- Sigui $v = x + \theta d \in P_e$; $0 < \theta \leq \theta^*$

$$c'v = c'x + \theta c'd = c'x + \theta [c_B', c_N'] \begin{bmatrix} -B^{-1} A_q \\ e_j \end{bmatrix} =$$

$$= c'x + \theta (\underbrace{c_q - c_B' \cdot B^{-1} \cdot A_q}_{r_q}) = c'x + \underbrace{\underbrace{\theta}_{>0} \cdot \underbrace{r_q}_{<0}}_{<0} \Rightarrow$$

$\Rightarrow c'v = c'x - \boxed{} \Rightarrow c'v < c'x \Rightarrow x$ NO SBF òptima

↓

CONTRADICCIÓ !! OK

✻ Interpretació del teorema b). Si x és SBF degenerada pot ser òptima i $r \not\geq [0]$.

Definició. Direcció de descens

Donat $(PL)e$, direm que la direcció factible d és de descens sobre $x \in P_e$ si $c'(x+\theta d) < c'x$, $\theta > 0$.

És a dir, és una DBF t.q. el cost reduït associat al vector q és < 0.

(d és de descens sobre $x \iff c'd < 0$)

* Si d és DBF sobre x SBF, sabem que $c'y = c'x + \theta^* c'd = c'x + \theta^* r_q$
  i) Els costos reduïts es poden expressar com $r_q = c'd$
  ii) Si P_e és NO degenerat ($\Rightarrow \theta^* > 0$), llavors la DBF d associada a $q \in \mathcal{N}$ és de descens $\iff r_q < 0$.

4.- Identificació de problemes (PL) il·limitats

Si tenim $d \geq 0$ DBF associada a la VNB x_q. Llavors:

$$d_B = -B^{-1} \cdot A_q \geq 0 \Rightarrow y_{B(i)} := \underbrace{x_{B(i)}}_{\geq 0} + \underbrace{\theta^*}_{\geq 0} \cdot \underbrace{d_{B(i)}}_{\geq 0} > 0 \quad \forall i = 1, \ldots, m$$

És a dir, $y = x + \theta d \geq 0 \Rightarrow y \in P_e \; \forall \theta > 0$.

↳ Conseqüències:
1) La longitud de pas θ^* NO està definida, ja que la trobàvem a partir de les $d_{B(i)} < 0$, i en aquest cas no n'hi ha cap. $\theta^* = \min_{i \in \emptyset} \left\{ \frac{-x_{B(i)}}{d_{B(i)}} \right\} = \text{"}+\infty\text{"}$ ($\nexists \theta^*$)

2) No tenim cap SBF adjacent. Al llarg de la semirecta $x + \theta d > 0$ (aresta del políedre) $\nexists$ SBF.

3) Si $r_q < 0$ (si trobem una DBF de descens), $z = c'x$ decreix sense límit al llarg de d (PL il·limitat)

$$z(x+\theta d) = z(\theta) = c'(x+\theta d) = c'x + \theta \overset{<0}{r_q} \xrightarrow{\theta \to +\infty} -\infty$$

5.- L'algorisme del símplex primal (ASP)

1. Inicialització: sigui (PL) $min\{c'x \mid x \in P_e\}$ P_e políedre estàndard de rang complet no buit no degenerat i la SBF inicial repressentada per $\mathcal{B}, \mathcal{N}, x_B$ i z

2. Identificació de SBF òptima i selecció de la variable no bàsica entrant q :

2.1 Es calculen els costos reduïts $r' = c_N' - c_B' B^{-1} A_N$

2.2. Si $r' \geq [0]$ llavors la s.b.f. actual és òptima: **STOP**.

Altrament, es selecciona una VNB q amb $r_q < 0$ (VNB entrant).

3. Càlcul de la DBF de descens. :

3.1. Es calcula $d_B = -B^{-1}A_q$ (DBF de descens associada a x_q)

3.2. Si $d_B \geq [0] \Rightarrow$ DBF de descens il·limitat: (PL) il.limitat: **STOP**

4. Càlcul de la passa màxima θ^* i selecció de la variable bàsica sortint $B(p)$:

4.1. Càlcul de la passa màxima al llarg de d_B : $\theta^* = \min_{\{i=1,\ldots,m \mid d_{B(i)}<0\}} \{-x_{B(i)}/d_{B(i)}\}$

4.2. Variable bàsica de sortida: $B(p)$ t.q. $\theta^* = -x_{B(p)}/d_{B(p)}$

5. Actualitzacions i canvi de base :

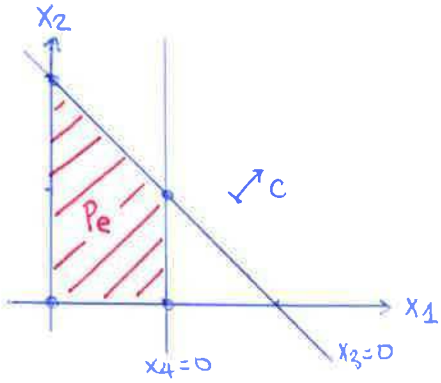
5.1. Actualització de les VB i f.o.: $x_B := x_B + \theta^* d_B$, $x_q := \theta^*$; $z := z + \theta^* r_q$

5.2. S'actualitzen $\mathcal{B} := \mathcal{B} \setminus \{B(p)\} \cup \{q\}$, $\mathcal{N} := \mathcal{N} \setminus \{q\} \cup \{B(p)\}$

6. Anada a 2.

Exemple.

$$(PL)_e \begin{cases} \min\ z = x_1 + x_2 \\ \text{s.a.:}\quad x_1 + x_2 + x_3 = 2 \\ \qquad\quad x_1 + x_4 = 1 \\ \qquad\quad x \geq 0 \end{cases}$$



1.- x^1: $\mathcal{B} = \{1,2\}$ $\quad \mathcal{N} = \{3,4\}$ $\qquad x_B = B^{-1} \cdot b$

$$x_B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{matrix} \to x_{B(1)} \\ \to x_{B(2)} \end{matrix}$$

$$z = c'_B \cdot x_B = \boxed{2}$$

- 1a iteració

2.- comencem el procés iteratiu → calcular costos reduïts

$$r' = \underbrace{c'_N}_{\text{vector de costos de les VNB}} - \underbrace{c'_B}_{\text{vector de costos de les VB}} \cdot B^{-1} \cdot \underbrace{A_N}_{\text{columnes de les VNB}} = (0,0) - (1,1) \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (\underbrace{-1}_{r_3}, \underbrace{0}_{r_4})$$

triarem la variable x_3 per a entrar a la base

$$c' = [\underbrace{1, 1}_{c_B}, \underbrace{0, 0}_{c_N}] \qquad (x_1\ x_2\ x_3\ x_4)$$

3.- $d_B = -B^{-1} \cdot A_q = -\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \not\geq 0$ $\Rightarrow$ Amb aquesta iteració no sabem si tenim un problema il·limitat
$q = 3$

$x_{B(1)}$ no té una component associada més petita que zero

$x_{B(2)}$ té una component associada < 0.

4.- $\theta^* = \min\limits_{i=2} \left\{ \dfrac{-x_{B(2)}}{d_{B(2)}} \right\} = \dfrac{-1}{-1} = \boxed{1}$ $\Rightarrow$ la variable que sortirà de la base és $B(2) = x_2$.

5.- $x^2 := x_B^1 + \theta^* \cdot d_B$

$$x^2 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{matrix} \leftarrow x_1 \\ \leftarrow x_2 \end{matrix}$$

ja que $y = x + \theta^* d$, i en la variable que entra: $x_q = 0$, $d_q = 1 \Rightarrow y_q = 0 + \theta^* \cdot 1 = \theta^*$

; $x_3 = \theta^* = 1$

Així doncs, $z = c'x \Rightarrow z^2 = z^1 + \theta^* \cdot r_q$

$$z^2 = 2 + 1 \cdot (-1)$$

$$z^2 = 2 + 1(-1) = \boxed{1}$$ → hem millorat la f. objectiu.

6.- x^2: $\mathcal{B} = \{1,3\}$ $\quad \mathcal{N} = \{2,4\}$

I anem al punt 2.-

• 2a iteració

2.- tenim $x_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow x_{B(1)} = x_1 \\ \leftarrow x_{B(2)} = x_3 \end{matrix}$ $\quad B = \begin{array}{c} \begin{matrix} x_1 & x_3 \end{matrix} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$ $\quad B^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$

ho hem calculat $x_B = B^{-1} \cdot b$

$$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [1, 0] - [1, 0] \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [\underset{r_2}{1}, \underset{r_4}{-1}]$$

Per tant triarem la variable x_4 per a entrar a la base.

3.- $d_B = -B^{-1} \cdot A_q = -\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \not\geq 0 \rightarrow$ no podem declarar el problema com a il·limitat

4.- $\theta^* = \min_{i=1} \left\{ \dfrac{-x_{B(1)}}{d_{B(1)}} \right\} = \dfrac{-1}{-1} = \boxed{1} \Rightarrow B(1) = x_1$ és la variable que sortirà de la base.

5.- $x^3 := x_B^2 + \theta^* \cdot d_B$ $\quad$ estem donant el valor de les variables en la nova base. x_1 passa a ser zero perquè surt de la base

$x^3 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \begin{matrix} \leftarrow x_1 \\ \leftarrow x_3 \end{matrix}$; $x_4 = \theta^* = 1$.

Així doncs, $z = c'x \Rightarrow z^3 = z^2 + \theta^* \cdot r_q$

$z^3 = 1 + 1(-1) = \boxed{0} \rightarrow$ millorem el valor de la f. objectiu.

6.- x^3 : $B = \{4, 3\}$ $\quad N = \{1, 2\}$

Anem al punt 2.-

• 3a iteració

2.- tenim $x_B = B^{-1} \cdot b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{matrix} \rightarrow x_{B(1)} = x_4 \\ \rightarrow x_{B(2)} = x_3 \end{matrix}$

$r' = c'_N - c'_B \cdot B^{-1} A_N$

$r' = [1, 1] - \underset{0}{[0, 0] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}} = [1, 1] \geq 0 \Rightarrow$ Hem arribat a l'òptim $= x^3 \Rightarrow B = \{4, 3\}$ $z = 0$.

6.- Implementació computacional del mètode del símplex. (MATLAB)

Funció simplexP_iter.m (MATLAB)

```
function [vb, vn, xb, z, iout] = simplexP_iter(c, A, b, vb, vn, xb, z)
:
:
if( rq >= 0)        iout = 2; return; end      % SBF òptima.
:
:
if (min(db) >=0 )  iout = 3; return; end      % Problema il·limitat.
:
:
if (theta == 0)     iout = 1; return; end      % SBF degenerada.
```

Funció simplexP.m

```
%
% Dades del problema
%
c=[-350, -300, 0, 0, 0]'  % Costos
A = [ 1,  1, 1, 0, 0;      % Matriu de coeficients
      9,  6, 0, 1, 0;
     12, 16, 0, 0, 1]
b= [ 200, 1566, 2880]'     % Vector de termes independents
%
vb=[3,4,5]                 % Conjunt inicial de VB
vn=[1,2]                   % Conjunt inicial de VNB
xb = A(:,vb)^(-1)*b;       % Valor inicial de les variables bàsiques
z = c(vb)'*xb;             % Valor inicial de la funció objectiu
%
iout=0;
niter = 0
while (iout == 0)
    niter = niter + 1;
    [vb, vn, xb, z, iout] = simplexP_iter( c, A, b, vb, vn, xb, z)
end
```

```
vb = 3 4 5
vn = 1 2
xb =
  200
 1566
 2880
z = 0
```

```
Iteració: 1
vb = 3 1 5
vn = 4 2
xb =
   26
  174
  792
z = -60900
iout = 0
```

```
Iteració: 2
vb = 2 1 5
vn = 4 3
xb =
   78.0000
  122.0000
  168.0000
z = -66100
iout = 0
```

```
Iteració: 3
vb = 2 1 5
vn = 4 3
xb =
   78.0000
  122.0000
  168.0000
z = -66100
iout = 2
```

7.- Fase I del símplex

Definició Fase I del símplex

Donat un problema de PL en forma estàndard: $(PL)\begin{cases} \min z = c'x \\ \text{s.a.: } Ax = b,\ b \geq 0 \\ x \geq 0 \end{cases}$

El seu problema associat de Fase I es defineix com: $(PL)_I \begin{cases} \min z_I = \sum_{i=1}^{m} y_i \\ \text{s.a.: } Ax + Iy = b \\ x, y \geq 0. \end{cases}$

La fase I del símplex és un procediment que permet trobar una SBF inicial definint un nou problema, que anomenem 'problema de Fase I', que té una SBF trivial.

Així doncs, la Fase I del símplex consisteix en la resolució del problema $(PL)_I$ amb l'algorisme del símplex a partir de: $\left.\begin{matrix} y = b \\ x = [0] \end{matrix}\right\}$ SBF trivial de $(PL)_I$

A partir d'aquí, considerem els possibles resultats que podem obtenir:

1.- $z^*_I > 0 \Rightarrow (PL)$ infactible

2.- $z^*_I = 0 \Rightarrow (PL)$ factible. Dos possibles casos:

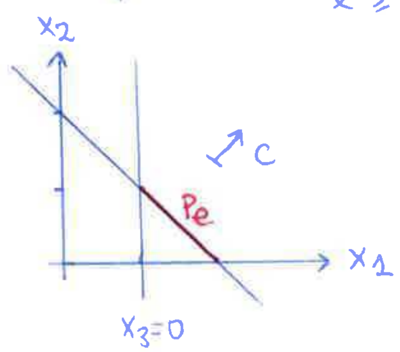
a) $\mathcal{B}^*_I$ no conté cap variable $y \Rightarrow \mathcal{B}^*_I$ és una SBF de (PL)

b) $\mathcal{B}^*_I$ conté alguna variable y:

- $y^*_B = [0] \Rightarrow \mathcal{B}^*_I$ és una SBF degenerada de $(PL)_I$
- Es pot obtenir una SBF de (PL) a partir de $\mathcal{B}^*_I$
  ↳ tindrem una SBF de (PL) degenerada

Exemple

$$(PL)\begin{cases} \min x_1 + x_2 \\ \text{s.a.: } x_1 + x_2 = 2 \\ \quad x_1 \geq 1 \\ \quad x \geq 0 \end{cases} \Rightarrow (PL)_e \begin{cases} \min x_1 + x_2 \\ \text{s.a.: } x_1 + x_2 = 2 \\ \quad x_1 - x_3 = 1 \\ \quad x \geq 0. \end{cases}$$



⚠ La única condició que imposem és que $b \geq [0]$
↳ vector dels termes independents.

Definim un nou problema → problema de fase I

$$(PL)_I = \begin{cases} \min z_I = x_4 + x_5 \\ \text{s.a.:} \quad x_1 + x_2 + x_4 = 2 \\ \qquad x_1 - x_3 + x_5 = 1 \\ \qquad x \geq 0 \end{cases}$$

Afegim una variable adicional a cada restricció
↓
creem una nova f. objectiu amb les variables artificials de Fase I.

Implementem l'ASP :

$B = \{4,5\} \quad N = \{1,2,3\}$ → vector de costos (de Fase I). $c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

$Ax + Iy = b \geq 0$ → Agafem com a base totes les variables artificials ⇒ y.

$B = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad x_B = B^{-1} \cdot b = b = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{matrix} \to x_4 \\ \to x_5 \end{matrix}$ } SBF de fase I

$z_I = x_4 + x_5 = 2 + 1 = \boxed{3}$

$N = \{1,2,3\}$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N$

$r' = [0,0,0] - [1,1] \cdot Id \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \dots = [-2,-1,1]$

$\begin{matrix} \downarrow & \downarrow & \downarrow \\ r_1 & r_2 & r_3 \end{matrix}$

$q = 1 \Rightarrow d_B = -B^{-1} \cdot A_q = -Id \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \ngeq 0 \Rightarrow$ No puc declarar el problema com a il·limitat.
↓
Entra la variable x_1

$\theta^* = \min_{i=4,5} \left\{ \frac{-x_{B(i)}}{d_{B(i)}} \right\} = \min \left\{ \frac{-2}{-1}, \frac{-1}{-1} \right\} = \boxed{1}$ → la variable bàsica que surt és $B(2) = x_5$

$x_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{matrix} \to x_4 \\ \to x_5 \end{matrix} \qquad x_1 = \theta^* = 1.$

Per tant, $z_I = z_I + \theta^* \cdot r_q = 3 + 1(-2) = \boxed{1}$

Comencem la 2a iteració:

$B = \{4,1\} \quad N = \{2,3,5\}$

$x_B = B^{-1} \cdot b = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{matrix} \to x_4 \\ \to x_1 \end{matrix}$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [0 \ 0 \ 1] - [1 \ 0] \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} =$

$= [-1 \ -1 \ 2]$ ⇒ Agafem $q = 2$ → x_2 és la variable que entra.

$\begin{matrix} \uparrow & \uparrow & \uparrow \\ r_2 & r_3 & r_5 \end{matrix}$

$d_B = -B^{-1}\cdot A_q = -\begin{pmatrix}1&-1\\0&1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix}\begin{matrix}\to d_4\\ \to d_1\end{matrix} \ngeq 0$ → no podem dir que el problema és il·limitat.

$\theta^* = \min_{i=4}\left\{\frac{-x_{B(i)}}{d_{B(i)}}\right\} = \frac{-1}{-1} = \boxed{1}$ → x_4 és la variable que surt.

$x_B = \begin{pmatrix}1\\1\end{pmatrix} + 1\begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}\begin{matrix}\to x_4\\ \to x_1\end{matrix}$; $x_2 = \theta^* = 1$

$z_I = z_I + \theta^* r_q = 1 + 1[-1] = \boxed{0}$

Comencem la 3a iteració:

$\mathcal{B} = \{2, 1\}$ $\quad \mathcal{N} = \{3, 4, 5\}$

$x_B = B^{-1}\cdot b = \begin{pmatrix}1&1\\1&0\end{pmatrix}^{-1}\cdot\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$

$r' = c'_N - c'_B\cdot B^{-1}\cdot A_n = [0\ 1\ 1] - [0\ 0]\cancel{\begin{bmatrix}0&1\\1&-1\end{bmatrix}\begin{bmatrix}0&1&0\\-1&0&1\end{bmatrix}} =$

$= [0\ 1\ 1]$ → hem trobat l'òptim

Per tant, la base de la fase I és $\mathcal{B} = \{2, 1\}$

↳ com que B^*_I no conté cap variable y tenim que B^*_I és SBF de (PL)e.

Així doncs hem trobat una SBF inicial per al nostre problema, ara el resoldrem aplicant l'ASP. (Fase II: aplicar l'algorisme del símplex sobre la base que hem trobat)

$\mathcal{B} = \{2, 1\}$ $\quad \mathcal{N} = \{3\}$

$x_B = B^{-1}\cdot b = \begin{pmatrix}1\\1\end{pmatrix}$ $\quad z = x_1 + x_2 = 1 + 1 = 2.$

$r' = c'_N - c'_B\cdot B^{-1}\cdot A_N = (0) - [1, 1]\begin{bmatrix}0&1\\1&-1\end{bmatrix}\begin{bmatrix}0\\-1\end{bmatrix} = \boxed{0}$ → hem trobat l'òptim.

8.- Convergència de l'ASP

Teorema⑥ Convergència de l'ASP.

Sigui (PL)e $\min\{c'x \mid x \in P_e\}$ problema de PL en forma estàndard, rang complet, factible sense cap SBF degenerada. Aleshores:

a) l'algorisme del símplex finalitza en un nombre finit d'iteracions.

b) l'algorisme del símplex finalitza en un dels següents estats:

i) Proporciona una SBF òptima.

ii) Identifica una SBF associada a una DBF de descens il·limitat → El problema (PL) és il·limitat

Demostració → immediata a partir de l'algorisme.

Així doncs, tenim la certesa que, quan no hi ha degeneració, en un nombre finit d'iteracions l'algorisme del símplex s'aturarà per un dels següents casos: i) $r' \geq [0]$ → ja hem trobat l'òptim

ii) $q : r_q < 0$, $d_B \geq [0]$ → problema il·limitat

↳ Què passa quan tenim degeneració?

- **Exemple de ciclat de l'algorisme del simplex amb degeneració:**

$$(P)\begin{cases} \min & -\frac{3}{4}x_1 + 20x_2 - \frac{1}{2}x_3 + 6x_4 \\ \text{s.a.:} & \frac{1}{4}x_1 - 8x_2 - 1x_3 + 9x_4 + x_5 = 0 \\ & \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 + x_6 = 0 \\ & x_3 + x_7 = 1 \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_6 \geq 0 \end{cases}$$

- **Aplicació de l'alg. del símplex amb selecció del cost reduït més negatiu:**

It. 1:	$\mathcal{B}=\{5,6,7\}$	$\mathcal{N}=\{1,2,3,4\}$	$x_B'=[0\ 0\ 1]$	$r'=[-3/4\ \ 20\ \ -1/2\ \ 6]$	$q=1$	$d_B'=[-1/4\ \ -1/2\ \ 0]$	$p=1$
It. 2:	$\mathcal{B}=\{1,6,7\}$	$\mathcal{N}=\{2,3,4,5\}$	$x_B'=[0\ 0\ 1]$	$r'=[-4\ \ -7/2\ \ 33\ \ 3]$	$q=2$	$d_B'=[32\ \ -4\ \ 0]$	$p=2$
It. 3:	$\mathcal{B}=\{1,2,7\}$	$\mathcal{N}=\{3,4,5,6\}$	$x_B'=[0\ 0\ 1]$	$r'=[-2\ \ 18\ \ 1\ \ 1]$	$q=3$	$d_B'=[-8\ \ -3/8\ \ -1]$	$p=1$
It. 4:	$\mathcal{B}=\{3,2,7\}$	$\mathcal{N}=\{1,4,5,6\}$	$x_B'=[0\ 0\ 1]$	$r'=[1/4\ \ -3\ \ -2\ \ 3]$	$q=4$	$d_B'=[21/2\ \ -3/16\ \ -21/2]$	$p=2$
It. 5:	$\mathcal{B}=\{3,4,7\}$	$\mathcal{N}=\{1,2,5,6\}$	$x_B'=[0\ 0\ 1]$	$r'=[-1/2\ \ 16\ \ -1\ \ 1]$	$q=5$	$d_B'=[-2\ \ -1/3\ \ 2]$	$p=1$
It. 6:	$\mathcal{B}=\{5,4,7\}$	$\mathcal{N}=\{1,2,3,6\}$	$x_B'=[0\ 0\ 1]$	$r'=[-7/4\ \ 44\ \ 1/2\ \ -2]$	$q=6$	$d_B'=[3\ \ -1/3\ \ 0]$	$p=2$
It. 7:	$\mathcal{B}=\{5,6,7\}$	$\mathcal{N}=\{1,2,3,4\}$	← mateixa base de la iteració 1 (CICLAT)				

Com a criteri per a escollir q (la variable que entrarà a la base) agafarem la VNB amb un cost reduït més negatiu.

Si entrem en un bucle, com en l'exemple d'aquí el costat, perdem la convergència de l'algorisme.

Això es pot evitar d'una manera "fàcil" amb el que anomenem criteris lexicogràfics de selecció de la variable d'entrada; és a dir, fixant unes regles específiques per tal de triar quina és la variable d'entrada i la variable de sortida. Hi ha diferents formes de fixar aquestes regles que eviten que perdem la convergència del símplex:

Regla de Bland (selecció del pivot de subíndex menor):

1.- Seleccionar com a VNB d'entrada la corresponent a l'índex menor de les que tinguin un cost reduït negatiu.

2.- Si en la selecció de la variable de sortida de la base es produeix un empat, seleccionar la variable bàsica amb índex menor.

Es pot demostrar que si s'aplica l'ASP amb la regla de Bland mai es produeix ciclat ⇒ l'algorisme del símplex acaba en un nombre finit d'iteracions.

9.- Complexitat algorísmica del símplex.

- Cost computacional del símplex:

$$\text{cost computacional} = \frac{\#\text{ operacions}}{\text{iteració}} \cdot \#\text{ iteracions}$$

- Nombre d'operacions/iteració:

a) $c'_B \cdot B^{-1}$: resolució sistema $B'\lambda = c_B \rightarrow O(m^2)$

b) $r' = c'_N - \lambda' A_N \rightarrow O(m) \sim O(m \cdot n)$

c) $d_B = -B^{-1} A_q$: resolució sistema $B\, d_B = -A_q \rightarrow O(m^2)$

↳ $O(m^2) \leq \#\text{ operacions/iteració} \leq O(m^2 + mn)$: cost polinòmic

- Nombre d'iteracions del símplex: polinòmic?

- Existeix un criteri de selecció de x_q t.q. el # iteracions ≤ polinomi en n i m?

↳ En la pràctica: s'observa que # iteracions $= O(m) \approx 3m$ (o $O(m \log n)$)

↳ En teoria: es poden trobar exemples on el # d'iteracions és exponencial.

Exemple. Problema de Klee-Minty (1972)

$$(P_{K\text{-}M}) \left\{ \begin{array}{ll} \min & -x_n \\ \text{s.a.:} & \varepsilon \leq x_1 \leq 1 \\ & \varepsilon(x_{i-1}) \leq x_i \leq 1 - \varepsilon(x_{i-1}) \quad i = 2, \ldots, n \\ & \varepsilon \in (0, 1/2) \end{array} \right.$$

Es pot demostrar que:

i) El políedre associat a $(P_{K\text{-}M})$ té 2^n vèrtexs (punts extrems).

ii) El símplex pot necessitar $2^n - 1$ iteracions (en el pitjor cas)

En conclusió, no se sap si # iteracions del símplex es pot expressar com a polinomi de n i m. És una qüestió fonamental de la matemàtica moderna.

TEMA 3: Teoria de dualitat

1.- Teoria de jocs (NO ENTRA)

Joc finit de suma zero amb dos jugadors. Cada jugador té un conjunt d'estratègies pures:
→ estratègies pures jugador 1: $J_1 = \{1, \dots, m\}$
↓ estratègies pures jugador 2: $J_2 = \{1, \dots, n\}$

A partir d'aquí podem desenvolupar una matriu de guany (J_1)/pèrdues (J_2) associades a les estratègies pures:

$$A = \begin{array}{c} \text{estr. } J_2 \\ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \end{array} \text{estr. } J_1$$

⊛ Si es produeix la jugada $(J_1, J_2) = (i, j) \Rightarrow$
⇒ el jugador 1 rep a_{ij} i el jugador 2 paga a_{ij} (joc de suma zero).

⊛ Estratègia mixta: distribució de probabilitat del conjunt d'estratègies pures (freqüència amb la que es jugarà cada estratègia)
- Jugador 1: $Y = \{y \in \mathbb{R}^m \mid \sum_{i=1}^{m} y_i = 1,\ 0 \le y_i \le 1,\ i = 1, \dots, m\}$
- Jugador 2: $X = \{x \in \mathbb{R}^m \mid \sum_{i=1}^{n} x_i = 1,\ 0 \le x_i \le 1,\ i = 1, \dots, n\}$

- **Jugada òptima jugador 1, criteri $\boldsymbol{maximin}$:**
  "El jugador 1 maximitza l'esperança matemàtica del seu guany mínim"

$$z_1^* = \max_y \left\{ z_1(y) = \min_{i=1,\dots,n} \left\{ \sum_{j=1}^{m} a_{ji} y_j \right\} \right\}$$

- **El problema de (PL) associat a la jugada òptima del jugador 1 és:**

$$(P_1) \begin{cases} \max\limits_{y, z_1} & z_1 & & \\ \text{s.a.:} & \sum_{j=1}^{m} a_{ji} y_j & \ge z_1 & i = 1, \dots, n \\ & \sum_{j=1}^{m} y_j & = 1 & \\ & y_j & \ge 0 & j = 1, \dots, m \end{cases}$$

- **Jugada òptima jugador 2, criteri $\boldsymbol{minimax}$:**
  "El jugador 2 minimitza l'esperança matemàtica de la seva pèrdua màxima"

$$z_2^* = \min_x \left\{ z_2(x) = \max_{j=1,\dots,m} \left\{ \sum_{i=1}^{n} a_{ji} x_i \right\} \right\}$$

- **El problema de (PL) associat a la jugada òptima del jugador 2 és:**

$$(P_2) \begin{cases} \min\limits_{x, z_2} & z_2 & & \\ \text{s.a.:} & \sum_{i=1}^{n} a_{ji} x_i & \le z_2 & j = 1, \dots, m \\ & \sum_{i=1}^{n} x_i & = 1 & \\ & x_i & \ge 0 & i = 1, \dots, n \end{cases}$$

- **Exemple: "pares o nones" amb dos dits**
  - Si la suma dels dits és senar, el jugador 1 rep del jugador 2 la suma dels dits en euros.
  - Si la suma dels dits és parell, el jugador 1 paga al jugador 2 la suma dels dits en euros.
  - Matriu de guanys J_1:

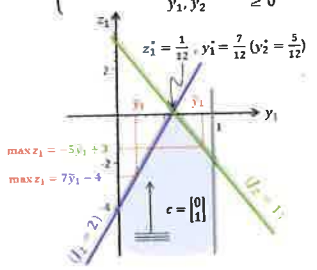
$$A = \begin{array}{cc} & J_2 \\ & \begin{array}{cc} 1 & 2 \end{array} \\ & \begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix} \end{array} \begin{array}{c} 1 \\ 2 \end{array} \Big\} J_1$$

  - Problema maximin jugador 1:

$$(P_1) \begin{cases} \max\limits_{y, z_1} & z_1 & \\ \text{s.a.:} & & \\ & -2y_1 + 3y_2 & \ge z_1 \\ & 3y_1 - 4y_2 & \ge z_1 \\ & y_1 + y_2 & = 1 \\ & y_1, y_2 & \ge 0 \end{cases}$$

- **Exemple: "pares o nones" amb dos dits**
  - Resolució del problema maximin jugador 1:

$$(P_1) \begin{cases} \max\limits_{y_1, y_2, z_1} & z_1 & \\ \text{s.a.:} & & \\ & -2y_1 + 3y_2 & \ge z_1 \\ & 3y_1 - 4y_2 & \ge z_1 \\ & y_1 + y_2 & = 1 \\ & y_1, y_2 & \ge 0 \end{cases} \xrightarrow{y_2 = 1 - y_1} (P_1) \begin{cases} \max\limits_{y_1, z_1} & z_1 & & \\ \text{s.a.:} & & & \\ & -5y_1 + 3 & \ge z_1 & (J_2 = 1) \\ & 7y_1 - 4 & \ge z_1 & (J_2 = 2) \\ & y_1 \in [0,1] & & \end{cases}$$



  - La recta $z_1 = -5y_1 + 3$ representa el valor esperat dels beneficis de J_1 en funció del valor de y_1 a les partides on J_2 juga l'estratègia 1.
  - La recta $z_1 = 7y_1 - 4$ representa el valor esperat dels beneficis de J_1 en funció del valor de y_1 en les partides on J_2 juga l'estratègia 2.
  - Per a cada valor de $y_1 \in [0,1]$:
    $\max z_1 = \min\{-5y_1 + 3, 7y_1 - 4\}$.
  - $y_1^* = \frac{7}{12}$ ($y_2^* = \frac{5}{12}$) és el valor de y_1 on el mínim entre les dues rectes és màxim ($z_1^* = \frac{1}{12}$).

- **Ta. Minimax (Ta. Principal de Ta. de Jocs, von Neumann 1928(1)) :**

Les estratègies òptimes y^* i x^* pels jugadors 1 i 2 existeixen i satisfan:

$$z_1^* = z_2^*.$$

2.- Definició i formulació del problema dual

Definició. Problema dual.

Sigui el problema de programació lineal (PL) $\min\{c'x \mid x \in P\}$ amb P poliedre. El problema dual (D) associat a (P) es defineix com el problema de programació lineal que s'obté a través de la següent taula de transformació:

	PROBLEMA PRIMAL (P)		PROBLEMA DUAL (D)	
Funció objectiu:	$\min c'x$	$\longleftrightarrow$	$\max \lambda' b$	: Funció objectiu
Constriccions primals, $j=1,\dots,m$:	$a_j x \geq b_j$	$\longleftrightarrow$	$\lambda_j \geq 0$	: Variables duals, $j=1,\dots,m$
	$a_j x \leq b_j$	$\longleftrightarrow$	$\lambda_j \leq 0$	
	$a_j x = b_j$	$\longleftrightarrow$	λ_j lliure	
Variables primals, $i=1,\dots,n$:	$x_i \geq 0$	$\longleftrightarrow$	$\lambda' A_i \leq c_i$	: Constriccions duals, $i=1,\dots,n$
	$x_i \leq 0$	$\longleftrightarrow$	$\lambda' A_i \geq c_i$	
	x_i lliure	$\longleftrightarrow$	$\lambda' A_i = c_i$	

Exemple.

$$(PL)\begin{cases} \min & x_1 + 2x_2 + 3x_3 \\ \text{s.a.:} & -x_1 + 3x_2 = 5 \\ & 2x_1 - x_2 + 3x_3 \geq 6 \\ & x_3 \leq 4 \\ & x_1 \geq 0,\ x_2 \leq 0,\ x_3 \text{ lliure} \end{cases} \rightarrow (D)\begin{cases} \max & 5\lambda_1 + 6\lambda_2 + 4\lambda_3 \\ \text{s.a.:} & -\lambda_1 + 2\lambda_2 \leq 1 \\ & 3\lambda_1 - \lambda_2 \geq 2 \\ & 3\lambda_2 + \lambda_3 = 3 \\ & \lambda_1 \text{ lliure},\ \lambda_2 \geq 0,\ \lambda_3 \leq 0 \end{cases}$$

* Hi ha tantes variables del problema dual com constriccions del problema primal i hi ha tantes constriccions en el problema dual com variables en el problema primal. De la mateixa manera trobem el signe de les variables i les constriccions del dual.

* El dual del dual és el primal.

3.- Teoremes de dualitat

Els teoremes de dualitat estudien les relacions entre les propietats dels problemes (PL) i (D).

En ocasions usarem el fet que el dual (D) d'un problema (PL) qualsevol i el dual $(D)_e$ de la seva forma estàndard $(PL)_e$ són equivalents.

Teorema ⑦. Equivalència duals forma estàndard.

Suposem que hem transformat un problema (PL) a la seva forma estàndard (PL)e de rang complet. Llavors els problemes duals de (PL) i (PL)e són equivalents en el sentit que o bé són tots dos infactibles o bé tenen el mateix cost òptim.

<u>Demostració</u> $\rightarrow$ la farem a problemes

Teorema ⑧. Teorema feble de dualitat.

Sigui x solució factible del problema (PL), i sigui λ solució factible del problema dual (D) associat. Llavors es satisfà:

$$\lambda' b \leq c' x$$

<u>Demostració</u>

Considerem un vector primal x i un vector dual λ qualssevol:

$x \in \mathbb{R}^n,\ \lambda \in \mathbb{R}^m$

$$\left.\begin{array}{ll} u_j = \lambda_j(a'_j x - b_j) & j = 1,\dots,m \\ v_i = (c_i - \lambda' A_i)x_i & i = 1,\dots,n \end{array}\right\} \text{si } x \text{ i } \lambda \text{ són factibles, aleshores } \begin{array}{ll} u_j \geq 0 & j = 1,\dots,m \\ v_i \geq 0 & i = 1,\dots,n \end{array}$$

- $u_j = \lambda_j(a'_j x - b_j)$ $\rightarrow$ tenim 3 casos:

1) $(a'_j x - b_j) = 0 \Rightarrow u_j = 0$. la constricció del prob. primal és de $\leq$

2) $(a'_j x - b_j) < 0 \Rightarrow a'_j x \leq b_j \Rightarrow \lambda_j \leq 0 \Rightarrow u_j \geq 0$

x sol. factible · la constricció del prob. primal és de $\geq$

3) $(a'_j x - b_j) > 0 \Rightarrow a'_j x \geq b_j \Rightarrow \lambda_j \geq 0 \Rightarrow u_j \geq 0$

- $v_i = (c_i - \lambda' A_i)x_i$ $\rightarrow$ tenim 3 casos:

$$\left.\begin{array}{l} 1)\ (\dots) = 0 \Rightarrow \dots \Rightarrow v_i = 0 \\ 2)\ (\dots) < 0 \Rightarrow \dots \Rightarrow v_i \geq 0 \\ 3)\ (\dots) > 0 \Rightarrow \dots \Rightarrow v_i \geq 0 \end{array}\right\} \text{Es fa de la mateixa manera que } u_j.$$

Així doncs tenim:

$$\left.\begin{array}{l} \sum_j u_j = \lambda' A x - \lambda' b \\ \sum_i v_i = c' x - \lambda' A x \end{array}\right\} \sum_j^{\textcircled{\scriptsize 1}} u_j + \sum_i^{\textcircled{\scriptsize 2}} v_i =$$

ho acabem de veure perquè x i λ són factibles.

$$= \cancel{\lambda' A x} - \lambda' b + c' x - \cancel{\lambda' A x} = c' x - \lambda' b \geq 0 \Rightarrow c' x \geq \lambda' b \ \underline{\text{OK}}$$

Gràficament: $$\underbrace{c' x}_{\substack{\text{S.F.} \\ \text{prob primal} \\ = \\ z_p}} \geq \underbrace{\lambda' b}_{\substack{\text{S.F.} \\ \text{prob. dual} \\ = \\ z_d}}$$

el valor de la f. obj associat a les solucions primals serà una fita superior del valor de la f. dual associat a les solucions duals.

Corol·laris:

i) Si (PL) és il·limitat, aleshores (D) és infactible.

Demostració: $\nexists \lambda \in \mathbb{R}^m \mid \lambda' b \leq -\infty$

ii) Si (D) és il·limitat, aleshores (PL) és infactible.

Demostració: $\nexists x \in \mathbb{R}^n \mid c'x \geq +\infty$

iii) Siguin x i λ solucions factibles (PL) i (D) respectivament tals que $\lambda' b = c'x$. Llavors x i λ són òptimes

Demostració: trivial.

Teorema 9 Teorema fort de dualitat

Si un problema de programació lineal (PL) té solució òptima, el seu dual (D) també en té, i els valors respectius de la funció objectiu coincideixen

Demostració

Estudiem la relació entre dues solucions òptimes.

1) Considerem (PL) en forma estàndard de rang complet i amb solució òptima (x)

↳ sabem que l'ASP + regla de Bland ⇒ trobarem una SBF òptima.

Veurem que $\lambda' = c'_B \cdot B^{-1}$ ($\lambda \in \mathbb{R}^m$) és i) factible dual

ii) $\lambda' b = c'x$

i) $r' = c'_N - \underbrace{c'_B \cdot B^{-1}}_{\lambda'} \cdot A_N = c'_N - \lambda' A_N \geq 0$

no ho entenc.

$= c'_N - \lambda' A_N \geq 0 \Rightarrow \lambda' A_N \leq c'_N$

Comprovem que $\lambda' A \leq c' \Rightarrow c'_B \cdot B^{-1}(B, A_N) =$

$\lambda' = c_B' \cdot B^{-1}$ $\quad A = [B, A_N]$

$= [c'_B, \underbrace{c'_B \cdot B^{-1}}_{\lambda'} \cdot A_N] \leq [c'_B, c'_N] = c'$ OK

corol·lari 8(iii)

ii) $\lambda' b = c'_B \cdot \underbrace{B^{-1} \cdot b}_{x_B} = c'_B \cdot x_B = c'x \Rightarrow \lambda$ i x són òptimes OK

↳ ja ho sabíem.

2) Què passa si (PL) té solució òptima i no està en forma estàndard?

↳ Ho podem resoldre fàcilment, ja que podem transformar el problema a forma estàndard. ($P \rightarrow P_e$) tindrà rg complet i la mateixa solució que P.

$z^*_{(P)} = z^*_{(Pe)} = z^*_{(D)} \Rightarrow z^*_{(P)} = z^*_{(D)}$, λ és òptim de (D).

$z^*_{(Pe)} = z^*_{(De)}$

Així doncs, si som capaços de trobar solució òptima del primal, sabem que el dual també té solució òptima i la podem trobar.

Corol·lari: Si (P) de rang complet té solució llavors (D) té solució i l'òptim dual és $\lambda' = c'_B B^{-1}$

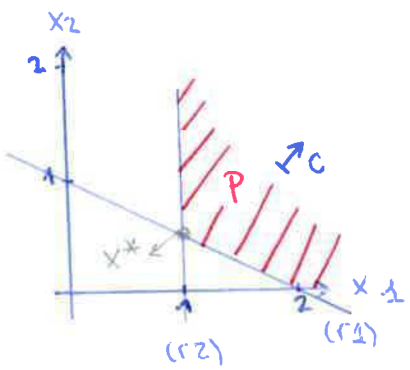
Demostració: immediata, a partir de la demostració del Teorema 9.

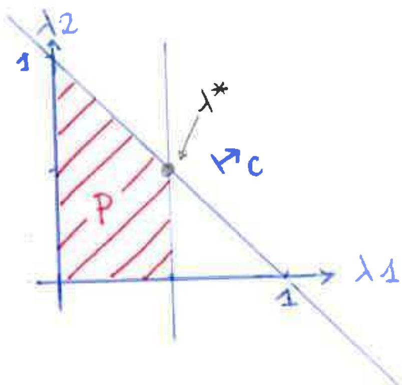
Possibles combinacions (P)–(D): els teoremes de dualitat fixen la següent relació de possibles casos:

(P) \ (D)	òptim	il·limitat	infactible
òptim	possible	impossible	impossible
il·limitat	impossible	impossible	possible
infactible	impossible	possible	possible

Exemple.

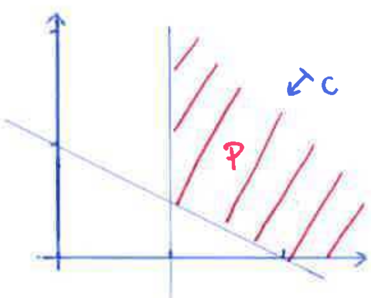
$$(P)\left\{\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a.:} & x_1 + 2x_2 \geq 2 \\ & x_1 \geq 1 \\ & x \geq 0 \end{array}\right. \rightarrow (D)\left\{\begin{array}{ll} \max & 2\lambda_1 + \lambda_2 \\ \text{s.a.:} & \lambda_1 + \lambda_2 \leq 1 \\ & 2\lambda_1 \leq 1 \\ & \lambda \geq 0 \end{array}\right.$$

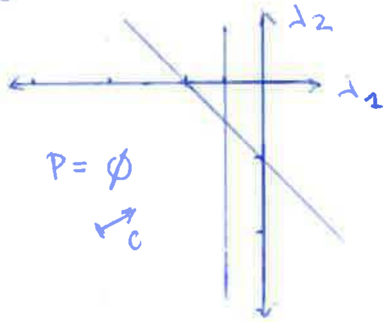




Exemple.

$$(P)\left\{\begin{array}{ll} \min & -x_1 - x_2 \\ \text{s.a.:} & x_1 + 2x_2 \geq 2 \\ & x_1 \geq 1 \\ & x \geq 0 \end{array}\right. \rightarrow (D)\left\{\begin{array}{ll} \max & 2\lambda_1 + \lambda_2 \\ \text{s.a.:} & \lambda_1 + \lambda_2 \leq -1 \\ & 2\lambda_1 \leq -1 \\ & \lambda \geq 0 \end{array}\right.$$





Teorema ⑨ [10] Teorema de folga complementària.

Siguin x i λ solucions factibles de (P) i (D) respectivament. Els vectors x i λ són solucions òptimes si i només si:

$$\left.\begin{array}{ll} \lambda_j(a'_j x - b_j) = 0 & j = 1,\dots,m \\ (c_i - \lambda' A_i)\, x_i = 0 & i = 1,\dots,n \end{array}\right\} \text{ condicions de folga complementària (CFC)}$$

Únic teorema que estableix unes condicions necessàries i suficients d'optimalitat

Demostració

$\Rightarrow$) x i λ òptimes $\Rightarrow$ CFC $= 0$

Siguin x i λ factibles $\Rightarrow$ CFC $\begin{cases} u_j \geq 0 & j = 1,\dots,m \\ v_i \geq 0 & i = 1,\dots,n \end{cases}$ OK $\rightarrow$ volem que siguin exactament zero

Prenem x i λ òptims $\overset{T9}{\Rightarrow}$ $c'x - \lambda'b = 0$

$$\sum_j \underbrace{u_j}_{\geq 0} + \sum_i \underbrace{v_i}_{\geq 0} = c'x - \lambda'b = 0 \Rightarrow \begin{cases} u_j = 0 & j = 1,\dots,m \\ v_i = 0 & i = 1,\dots,m \end{cases}$$ OK

$\Leftarrow$) $\sum_j u_j + \sum_i v_i \underset{\forall x \text{ i } \lambda \text{ factibles}}{=} c'x - \lambda'b \underset{\text{CFC}}{=} 0 \overset{\text{corol·lari T8}}{\Rightarrow} x$ i λ òptims

4.- Algorisme del símplex dual.

Definició. Solució bàsica factible dual.

Sigui el problema de programació lineal (PL) en forma estàndard.

Una solució bàsica factible dual és tota solució bàsica de (PL) amb $r \geq [0]$

↳ És a dir, són les solucions bàsiques primals on el vector de ~~que tenen~~ costos reduïts de les variables no bàsiques ≥ 0.

⊛ Si $r \geq 0$, llavors pel el teorema fort de dualitat sabem que $\lambda' = c'_B \cdot B^{-1}$ és una solució factible pel problema dual (D).

Una solució bàsica factible dual pot no ser factible primal

↳ una solució factible dual i factible primal és òptima.

- Factibilitat primal: $x_B = B^{-1}b \geq 0$
- Factibilitat dual: $r \geq 0$

Definició. Algorisme del símplex dual (ASD)

L'algorisme del símplex dual és un algorisme que permet resoldre problemes de PL en forma estàndard a partir de solucions bàsiques factibles duals basant-se en la següent estratègia:

a) Es determina si la SBFD actual és factible primal ⇒ òptima

b) Si la SBFD actual no és factible primal es troba, si existeix, una SBFD adjacent a l'actual que minori el valor de la f. objectiu dual, i es pren aquesta com a nova solució bàsica actual.

❀ Interès del símplex dual:

situacions on es disposa d'una SBFD infactible primal.

↳ Anàlisi de sensibilitat: canvis en A i/o B.

↳ Programació lineal entera (algorisme de Branch & Bound)

Forma estàndard (NO ENTRA)

Idea: aplicarem el mateix desenvolupament del símplex primal a la forma estàndard modificada del problema (D).

$$(P)e\left\{\begin{array}{l}\min z_P = c'x\\ \text{s.a.: } Ax = b\\ \quad x \geq 0\end{array}\right. \rightarrow (D)\left\{\begin{array}{l}\max z_D = \lambda' b\\ \text{s.a.: } A'\lambda \leq c\\ \quad \lambda \text{ lliure.}\end{array}\right. \rightarrow$$

afegim variables per a que es compleixi la igualtat.

una a cada constricció per això fem $I_n \cdot r$

$$\rightarrow (D)\left\{\begin{array}{l}\min -z_D = -\lambda' b\\ \text{s.a.: } A'\lambda + I_n r = c\\ \quad r \geq 0\\ \quad \lambda \text{ lliure}\end{array}\right. \rightarrow (D)\left\{\begin{array}{l}\min -z_D = -b'\lambda\\ \text{s.a.: } [A'\ I_n]\begin{bmatrix}\lambda\\ r\end{bmatrix} = c\\ \quad r \geq 0\\ \quad (\lambda \text{ lliure})\end{array}\right.$$

Els costos reduïts del primal són les folgues del dual

Considerem ara la base $\mathcal{B}$ associada a una SBFD de (P)e

$$\mathcal{B} \text{ SBFD} \Rightarrow r'_N = c'_N - c'_B \cdot B^{-1} \cdot A_N \geq 0$$

(costos reduïts ≥ 0)

$$[A'\ I_n]\begin{bmatrix}\lambda\\ r\end{bmatrix} = \begin{array}{c}\mathcal{B}\rightarrow\\ \mathcal{N}\rightarrow\end{array}\begin{bmatrix}B' & 0 & I_m\\ A'_N & I_{n-m} & 0\end{bmatrix}\begin{bmatrix}\lambda\\ r_N\\ r_B\end{bmatrix} = \underbrace{\begin{bmatrix}B' & 0\\ A'_N & I_{n-m}\end{bmatrix}}_{B_D}\underbrace{\begin{bmatrix}\lambda\\ r_N\end{bmatrix}}_{y_B} + \underbrace{\begin{bmatrix}I_m\\ 0\end{bmatrix} r_B}_{y_n} =$$

$$= \begin{bmatrix}c_B\\ c_N\end{bmatrix}$$

→ això és la partició de les constriccions duals induïda per la base primal $\mathcal{B}$.

Definició. Solució dual y associada a una SBFD de $(P)e$:

És la solució de $[A' \; I_n]\begin{bmatrix}\lambda \\ r\end{bmatrix} = \begin{bmatrix} c_B \\ c_N \end{bmatrix}$ amb $r_B = [0]$

$$y_B = \begin{bmatrix}\lambda \\ r_N\end{bmatrix} \in \mathbb{R}^n, \qquad y_N = r_B = [0] \in \mathbb{R}^m$$

Veurem que $y' = [y'_B \;\; y'_N]$ és una SBF del políedre dual en forma estàndard:

$$D_e = \left\{ y = \begin{bmatrix}\lambda \\ r\end{bmatrix} \in \mathbb{R}^{n+m} \;\middle|\; [A' \;\; I_n]\begin{bmatrix}\lambda \\ r\end{bmatrix} = c,\; r \geq 0 \right\}$$

Com que D_e està en forma estàndard modificada, necessitem una definició alternativa de SBF.

Definició. def. alternativa de SBF.

El vector $x \in P \subset \mathbb{R}^n$ és una solució bàsica factible del políedre $P \iff$ $\iff$ hi ha com a mínim n constriccions actives linealment independents sobre x.

La SBF serà degenerada si hi ha més de n constriccions actives linealment independents sobre x.

Proposició. La solució dual y associada a una SBFD de $(P)e$ és una SBF del políedre dual D_e amb matriu bàsica $B_D = \begin{bmatrix} B' & 0 \\ A_N' & I_{n-m}\end{bmatrix}$.

Costos reduïts duals

Es tracta ara de reproduir la passa del símplex primal aplicada a la resolució del problema dual D_e no degenerat a partir de la SBF de $(D)e$ $y' = [\lambda' \;\; r'_N \;\; 0]$:

$$y'_B = [\lambda' \;\; r'_N], \quad b'_B = [-b' \;\; 0], \quad B_D^{-1} = \begin{bmatrix} \overset{(B^{-1})^T}{B^{-T}} & 0 \\ -A'_N B^{-T} & I_{n-m} \end{bmatrix}$$

$$y_N = r_B, \quad b'_N = [0]$$

Costos reduïts duals: $r_D = b'_N - b'_D \cdot B_D^{-1}\begin{bmatrix} I_m \\ 0 \end{bmatrix} = [0] - [-b' \;\; 0]\begin{bmatrix} B^{-T} & 0 \\ -A'_N B^{-T} & I_{n-m}\end{bmatrix}\begin{bmatrix} I_m \\ 0 \end{bmatrix} =$

$$= [b' \cdot B^{-T} \;\; 0]\begin{bmatrix} I_m \\ 0 \end{bmatrix} = b' B^{-T} = x'_B$$

$\Downarrow$

La condició d'optimalitat de $(D)e$ és que $x_B \geq [0]$ ($\equiv$ factible primal).

Direccions bàsiques factibles duals

- DBF de $(D)_e$, d_{B_D}:

Sigui $\boxed{r_{B(p)} \text{ variable no bàsica entrant}}$ amb $x_{B(p)} < 0$.

$$d_{B_D} = \begin{bmatrix} d_\lambda \\ d_{r_N} \end{bmatrix} = -B_D^{-1} \begin{bmatrix} e_p \\ 0 \end{bmatrix} = \begin{bmatrix} -B^{-T} & 0 \\ A'_N B^{-T} & -I_{n-m} \end{bmatrix} \begin{bmatrix} e_p \\ 0 \end{bmatrix} = \begin{bmatrix} -B^{-T} \\ A'_N B^{-T} \end{bmatrix} e_p$$

si indiquem per β_p la fila p-èsima de B^{-1}, $\beta_p = e'_p \cdot B^{-1}$:

$$d_{B_D} = \begin{bmatrix} d_\lambda \\ d_{r_N} \end{bmatrix} = \begin{bmatrix} -B^{-T} e_p \\ A'_N B^{-T} e_p \end{bmatrix} = \begin{bmatrix} -\beta'_p \\ A'_N \beta'_p \end{bmatrix} \Rightarrow \begin{array}{l} d_\lambda = -\beta'_p \\ d_{r_N} = (\beta_p A_N)' \end{array}$$

$$\downarrow$$

d_{B_D} és direcció de descens: $[-b' \;\; 0]\, d_{B_D} = x_{B(p)} < 0$

- Longitud de pas dual θ^*_D: passa màxima que conserva la factibilitat dual

$$\begin{bmatrix} \lambda \\ r_N \end{bmatrix} := \begin{bmatrix} \lambda \\ r_N \end{bmatrix} + \theta^*_D \begin{bmatrix} d_\lambda \\ d_{r_N} \end{bmatrix} = \begin{bmatrix} \lambda + \theta^*_D d_\lambda \\ r_N + \theta^*_D d_{r_N} \end{bmatrix} \begin{array}{l} \lessgtr 0 \\ \geq 0 \end{array} \Bigg\} \Rightarrow$$

$$\Rightarrow \theta^*_D = \min_{\{j \in \mathcal{N} \mid d_{r_{N_j}} < 0\}} \left\{ \frac{-r_j}{d_{r_{N_j}}} \right\} = \frac{-r_q}{d_{r_{N_q}}}$$

- Problema il·limitat $(D)_e$: $d_{r_N} \geq 0 \Rightarrow (D)_e$ il·limitat

Algorisme del símplex dual (ASD)

1. Sigui la SBFD $\mathcal{B}$ amb valors: $B, x_B, r, c_B, c_N, A_N, z$.

2. Identificació de SBF òptima i selecció de la variable bàsica sortint $B(p)$:

2.1. Si $x_B \geq [0]$ llavors la SBF actual és òptima: **STOP**.

Altrament, es selecciona una VB p amb $x_{B(p)} < 0$ (VB sortint).

3. Càlcul de la DBF de $(D)_e$:

3.1. Es calcula $d_{r_N} = (\beta_p A_N)'$ (β_p: fila p-èssima de B^{-1})

3.2. Si $d_{r_N} \geq [0]$ llavors problema $(D)_e$ il.limitat ($\Rightarrow (P)_e$ infactible): **STOP**

4. Selecció de la variable no bàsica entrant q:

4.1. Càlcul de $\theta^*_D = \min_{\{j \in \mathcal{N} \mid d_{r_{N_j}} < 0\}} \left\{ \frac{-r_j}{d_{r_{N_j}}} \right\} = \frac{-r_q}{d_{r_{N_q}}}$. Es selecciona q com a VNB entrant.

5. Canvi de base i actualitzacions :

5.2. Act. variables duals: $r_N := r_N + \theta^*_D d_{r_N}$, $\lambda := \lambda - \theta^*_D \beta'_p$, $r_{B(p)} := \theta^*_D$; $z := z - \theta^*_D x_{B(p)}$

5.1. Act. variables primals: $d_B = -B^{-1} A_q$, $\theta^* = -\frac{x_{B(p)}}{d_{B(p)}}$, $x_B := x_B + \theta^* d_B$, $x_q := \theta^*$

5.2. Act. base: $\mathcal{B} := \mathcal{B} \setminus \{B(p)\} \cup \{q\}$, $\mathcal{N} := \mathcal{N} \setminus \{q\} \cup \{B(p)\}$.

6. Anada a 2.

Exemple. SB inicial associada a $x' = [0\ \ 0]$

$$(P)\begin{cases} \min\ z = x_1 + x_2 \\ \text{s.a.:}\ \ x_1 + 2x_2 \geq 2 \\ \quad\ \ x_1 \geq 1 \\ \quad\ \ x \geq 0 \end{cases} \rightarrow (P)e \begin{cases} \min\ z = x_1 + x_2 \\ \text{s.a.:}\ \ x_1 + 2x_2 - x_3 = 2 \\ \quad\ \ x_1 - x_4 = 1 \\ \quad\ \ x \geq 0 \end{cases}$$

volem que $x' = [0\ \ 0\ \ x_3\ \ x_4] \Rightarrow x' = [0,\ 0,\ -2,\ -1] \rightarrow$ infactible primal.

Aplicarem ASD
↳ No cal tenir el problema dual formulat

$\beta = \{3, 4\}$ $\quad B^{-1} = -Id_2$ $\quad x_B = B^{-1}\cdot b = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$

$N = \{1, 2\}$ $\quad A_N = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ $\quad c_N = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda' = c'_B \cdot B^{-1} = [0\ \ 0](-Id) = [0]$

$r = c'_N - c'_B \cdot B^{-1} \cdot A_N = c'_N - \lambda' A_N = [1\ \ 1] - [0] A_N = [1\ \ 1] \geq [0]$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \uparrow\ \ \uparrow$
$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad r_1\ \ r_2$

Comencem la iteració:

$\beta = \{3, 4\}$ $\quad N = \{1, 2\}$

- Identificació de SBF òptima i selecció de la VB sortint $B(p)$:

$x_B' = [-2\ \ -1] \ngeq [0] \Rightarrow$ no és SBF òptima.

$\quad\ \downarrow\quad\ \downarrow$
$\quad\ x_3\quad x_4$ $\quad\quad$ $B(1) = x_3$ variable sortint.

- Identificació de problema (D) il·limitat:

$\beta_1 = e'_1 B^{-1} = [1\ \ 0]\cdot(-Id) = [-1\ \ 0]$

↳ e_p és una columna on tot són zeros menys a la posició p, que hi ha un 1 → → la variable p entra a la base i per definició $d_p = 1$.

$d'_{r_N} = \beta_1 \cdot A_N = [-1\ \ 0]\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} = [-1\ \ -2] \ngeq [0] \rightarrow$ no es podem declarar il·limitat

- Selecció de la VNB entrant q:

$$\theta^* = \min_{\{j\in N \mid d_{r_{N_j}} < 0\}} \left\{ \frac{-r_j}{d_{r_{N_j}}} \right\} = \min\left\{ \frac{-1}{-1}, \frac{-1}{-2} \right\} = \frac{1}{2} \Rightarrow q = 2 \rightarrow \text{Entra } x_2$$

En la nostra actualització $B(1) = x_3$ sortirà de la base i x_2 entrarà.

↳ $r_N = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} := r_N + \theta^*_D\, d_{r_N} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}$ $\quad\Big|\quad r_{B(1)} = r_3 := \theta^*_D = \frac{1}{2}$

↳ $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} := \lambda - \theta^*_D\, \beta'_p = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 0 \end{bmatrix}$

↳ $z := z - \underbrace{\theta^*_D\, x_{B(1)}}_{1/2 \cdot x_3} = 0 - \frac{1}{2}(-2) = 1$

$\hookrightarrow d_B = -B^{-1} \cdot A_2^q = -\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \begin{matrix} \leftarrow x_3 \\ \leftarrow x_4 \end{matrix}$ | $\theta^* = -\frac{x_{B(1)}}{d_{B(1)}} = 1$

$\hookrightarrow x_B = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} := x_B + \theta^* d_B = \begin{bmatrix} -2 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ $\quad x_2 := \theta^* = 1$

$\hookrightarrow \mathcal{B} = \{2, 4\} \quad \mathcal{N} = \{1, 3\}$

$B^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & -1 \end{bmatrix} \quad x_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad r_N = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \quad \lambda = \begin{bmatrix} -1/2 \\ 0 \end{bmatrix}$

Segona iteració:

$\mathcal{B} = \{2, 4\} \quad \mathcal{N} = \{1, 3\}$

- Identificació de SBF òptima i selecció de la VB sortint $B(p)$:

$x_B = [\underset{\downarrow\, x_2}{1} \quad \underset{\downarrow\, x_4}{-1}] \not\geq [0] \Rightarrow$ no és l'òptim

$p = 2 \rightarrow B(2) = x_4$ VB sortint

- Identificació de problema (D) il·limitat:

$\beta_2 = e'_2 B^{-1} = [0 \quad 1] \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} = [0 \quad -1]$

$d'_{r_N} = \beta_2 A_N = \beta_2 A_N = [0 \quad -1] \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = [-1 \quad 0] \not\geq [0]$ no podem declarar el problema com a il·lim.

- Selecció de la VNB entrant q:

$\theta^* = \min_{\{j \in \mathcal{N} \mid d_{r_{N_j}} < 0\}} \left\{ \frac{-r_j}{d_{r_{N_j}}} \right\} = \min \left\{ \frac{-1/2}{-1} \right\} = \frac{1}{2} \Rightarrow q = 1 \rightarrow$ Entra x_1

- Canvi de base i actualitzacions:

$\hookrightarrow r_N = \begin{bmatrix} r_1 \\ r_3 \end{bmatrix} = r_N + \theta^*_D d_{r_N} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$ | $r_{B(2)} = r_4 := \theta^*_D = \frac{1}{2}$

$\hookrightarrow \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} := \lambda - \theta^*_D \beta_p = \begin{bmatrix} -1/2 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}$

$\hookrightarrow z := z - \theta^*_D x_{B(2)} = 1 - \frac{1}{2}(-1) = \frac{3}{2}$

$\hookrightarrow d_B = -B^{-1} \cdot A_1 = -\begin{pmatrix} 1/2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}$ | $\theta^* = -\frac{x_{B(2)}}{d_{B(2)}} = -\frac{-1}{1} = 1$

$\hookrightarrow x_B = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} := x_B + \theta^* \cdot d_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix} \quad x_1 = \theta^* = 1$

$\hookrightarrow \mathcal{B} = \{2, 1\} \quad \mathcal{N} = \{3, 4\}$

$B^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 0 & 1 \end{bmatrix} \quad x_B = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} \quad r_N = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \quad \lambda = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}$

Fem la tercera iteració:

$\beta = \{2,1\}$ $\quad N = \{3,4\}$

- Identificació de SBF òptima i selecció de la VB sortint $B(p)$:

$x_B = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} \geq [0] \Rightarrow$ SBF primal i dual

↳ hem vist que $r \geq [0]$

(a més l'ASD manté la factibilitat dual)

Per tant no seguim iterant. Solució òptima: $\beta^* = \{2,1\}$ $\quad x_B^* = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$

$N^* = \{3,4\}$ $\quad \boxed{z^* = 3/2}$

5.- Símplex dual: degeneració i convergència

Definició. SBFD de (P)e degenerada dual:

Es diu que una SBFD de (P)e és degenerada dual si $\exists j \in N$ tq $r_j = 0$

Propietats.

i) La base òptima de (P)e és degenerada dual $\Leftrightarrow$ (P)e té òptims alternatius

ii) Sigui β SBFD de (P)e. Aleshores β és degenerada dual $\Leftrightarrow$ la SBF de (D)e associada a β és SBF degenerada.

↳ una base primal degenerada dual és una base dual degenerada primal (del dual)

És a dir: Base del problema Primal: $r_j = 0$ $\quad j \in N \Rightarrow$ base del primal degenerada dual

Base del problema Dual: r_j és VB $\Rightarrow$ VB del dual $= 0 \Rightarrow$ degeneració primal en el dual

Teorema (15) Convergència de l'algorisme del símplex dual.

Si el problema (P)e no té cap SBFD amb degeneració dual, aleshores l'algorisme del símplex dual convergeix en un nombre finit d'iteracions.

Demostració

Cada iteració augmenta estrictament el valor de la funció dual $\lambda' b \Rightarrow$

$\Rightarrow$ no es repeteix cap SBFD de (P)e i el nombre de SBFD és finit.

✱ Si (P)e té SBFD degenerades, la regla de Bland (entre d'altres) assegura la convergència.

TEMA 4 : Programació Lineal Entera

1.- Conceptes bàsics

Definició Problema de programació lineal entera (PLE)

Quan una o diverses variables x_i, $i \in J$ d'un problema de PL només pot adoptar valors enters, es té un problema de Programació Lineal Entera.

* Els problemes de PLE són habituals quan les solucions fraccionals no tenen sentit. Les variables enteres també ens ajuden a construir models més acurats per a un nombre de problemes de presa de decisions.

Exemples. (NO ENTRA)

1.- Planificació de plantilles laborals

↳ L'objectiu és obtenir quants treballadors contractar a cada torn de forma que es minimitzin els costos de personal tot satisfent les necessitats de treballadors de cada dia.

- Formulació genèrica:
  - Paràmetres:
    - n : # de torns
    - c_i : salari torn i, $i = 1, \dots, n$
    - D_i : dies de descans torn i, $i = 1, \dots, n$
    - b_j : # de treballadors necessaris cada dia j, $j = 1, \dots, 7$
  - Variables de decisió:
    - x_i = # de treballadors assignats al torn i, $i = 1, \dots, n$
  - Formulació:

$$
(PE) \left\{ \begin{array}{ll}
\min\limits_{x} \; z = \sum_{i=1}^{n} c_i x_i & \rightarrow \text{Es minimitza el cost salarial} \\
\text{s.a.:} \; \sum_{i \mid j \notin D_i} x_i \geq b_j, \; j = 1, \dots, 7 & \rightarrow \text{Es satisfan les necessitats laborals} \\
x_i \geq 0, \; x_i \in \mathbb{Z}, \; i = 1, \dots, n &
\end{array} \right.
$$

2.- Selecció de projectes. Variables de decisió

- Formulació genèrica:
  - Paràmetres:

    n: # projectes

    m: # anys

    c_i: benefici esperat del projecte i, $i=1,\dots,n$

    a_{ij}: capital necessari projecte i any j; $i=1,\dots,n$ $j=1,\dots,m$

    b_j: capital total disponible any j, $j=1,\dots,m$

  - Variables de decisió:

$$x_i = \begin{cases} 1, & \text{el projecte es selecciona} \\ 0, & \text{el projecte NO es selecciona} \end{cases} ; \quad i=1,\dots,n$$ → variables binàries

  - Formulació:

$$(PE) \begin{cases} \max\limits_{x} \ z = \sum_{i=1}^{n} c_i x_i \\ \text{s.a.:} \ \sum_{i=1}^{n} a_{ij} x_i \le b_j \quad j=1,\dots,m \\ x \in \{0,1\}^n \end{cases}$$

→ Es maximitza el valor present net (NPV)

→ sense superar el capital disponible

✲ Les variables binàries són útils per a modelitzar condicions lògiques

3.- Problemes de producció amb costos fixos

↳ L'objectiu és obtenir el programa de producció que maximitzi el guany net (benefici menys costos) tot satisfent la disponibilitat de recursos.

- Formulació genèrica:
  - Paràmetres:

    n: # productes

    m: # processos

    c_i: benefici unitari producte i, $i=1,\dots,n$

    a_{ij}: hores consumides producte i procés j; $i=1,\dots,n$ $j=1,\dots,m$

    b_j: hores totals disponibles procés j, $j=1,\dots,m$

    k_i: costos configuració producte i, $i=1,\dots,n$

  - Variables de decisió:

    x_i: quantitat producte i, $i=1,\dots,n$

$$y_i = \begin{cases} 1, & x_i > 0 \\ 0, & x_i = 0 \end{cases} , \quad i=1,\dots,n$$

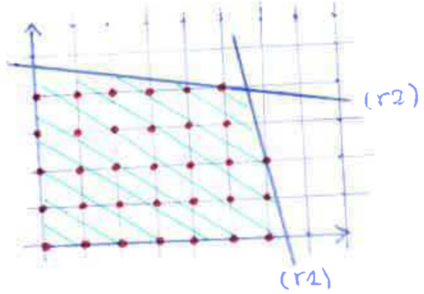
- Formulació:

$$(PE)\begin{cases} \max_{x,y} \; z = \sum_{i=1}^{n} (c_i x_i - k_i y_i) & \rightarrow \text{Es maximitza el benefici net} \\ \text{s.a.:} \; \sum_{i=1}^{n} a_{ij} x_i \leq b_j \quad j=1,\dots,m & \rightarrow \text{Disponibilitat de recursos} \\ x \leq M_i y_i \quad i=1,\dots,n & \rightarrow \text{Acoblament } x_i - y_i \\ x \geq 0 \quad y \in \{0,1\}^n \end{cases}$$

2.- Relaxació lineal

Definició. Relaxació lineal

Sigui PLE un problema de programació lineal entera, si en fem la representació gràfica:



K_{PLE}: regió factible del problema = {•}

↓

no és convexa ⇒ no podem utilitzar les propietats dels problemes de PL. →

→ tot i així, els problemes de PL juguen un paper molt important en la resolució de problemes PLE.

La relaxació lineal del problema PLE és un problema de PL que coincideix amb el problema de PLE en totes les seves característiques tret de la integritat de les seves variables. És a dir, eliminem la condició d'integritat de les variables, de manera que ens queda un problema PL amb una regió factible convexa que podem resoldre amb el temari dels capítols anteriors.

K_{RL}: regió factible de la relaxació lineal =

És fàcil veure que $K_{PLE} \subseteq K_{RL}$ → aquesta relació és la base per a resoldre els problemes PLE.

✲ La solució òptima de (RL) serà sempre millor o igual que la solució òptima de (PLE), si existeix solució òptima.

↳ $K_{PLE} \subseteq K_{RL}$ ⇒ la solució òptima de la relaxació lineal z^*_{RL} proporciona una fita del valor òptim de la f. objectiu del problema PLE z^*_{PE}

- Per a problemes de maximització el valor òptim de la RL és una fita superior del valor òptim de la f. objectiu del problema PLE
- Per a problemes de minimització el valor òptim de la RL és una fita inferior del valor òptim de la f. objectiu del problema PLE.

Relació solucions (PLE) - (RL)

(RL)

(PLE)	Sol. òptima	Infact.	Il·limitat
Sol. òptima	sí (1)	no	sí (4) ← !
Infactible	sí (2)	sí	sí (5)
Il·limitat	no (3)	no	sí

(1) Si $x^*_{RL} \in K_{PLE} \Rightarrow x^*_{RL} = x^*_{PLE}$ → la solució del problema enter coincideix amb la sol. del problema relaxat

(2) (PLE) $\min_{x \in \mathbb{Z}} \{ x_1 \mid \frac{1}{4} \leq x_1 \leq \frac{3}{4} \}$: (PLE) infactible, $x^*_{RL} = [\frac{1}{4}, x_2]'$

x_1 no pot prendre valors enters

(3) $z^*_{RL} \leq z^*_{PLE}$ ↳ la RL és una fita inferior del valor de la f obj del problema PLE

si $\exists\, z^*_{RL}$, z^*_{RL} és la fita inferior ⇒ z^*_{PLE} no és il·limitat

(4) (PLE) $\min_{x \in \mathbb{Z}^4} \{ -x_1 \mid x_3 - \sqrt{2}(x_1 - x_2) = 0,\ x_2 + x_4 = 1,\ x \geq 0 \}$

⚠ coeficient irracional

⚠

- (RL) il·limitat ; $x = [\alpha \ 0 \ \alpha\sqrt{2} \ 1]'$ factible $\forall \alpha \geq 0$.
- (PLE) òptim : $K_{PLE} = \{[0\ 0\ 0\ 1]', [1\ 1\ 0\ 0]'\}$ amb $x^*_{PLE} = [1\ 0\ 0\ 0]'$

$x_2 = 0$ $\boxed{x_4 = 1}$ $\boxed{x_2 = 1}$ $x_4 = 0$

$x_3 = \sqrt{2}(x_1 - 1) \Rightarrow \boxed{x_1 = 1}$, $x_3 = 0$

Si un prob. relaxat NO té sol. esperem que el seu problema original tampoc en tingui
↳ hem d'imposar una condició addicional

(5) (PLE) $\min_{x \in \mathbb{Z}^2} \{ x_1 + x_2 \mid \frac{1}{4} \leq x_1 \leq \frac{3}{4} \}$: (PLE) infactible, (RL) il·limitat.

Teorema (16). Sigui (PLE) $\min_{x \in \mathbb{Z}^n} \{ c'x \mid Ax = b,\ x \geq 0 \}$ i (RL) la seva relaxació lineal.
Si A té coeficients racionals ($a_{ji} \in \mathbb{Q}$) i (RL) és il·limitat llavors el problema (PLE) és o bé il·limitat o bé infactible.

↳ Per tant, el problema enter té solució òptima només si el problema relaxat en té.

(Nosaltres assumirem que A racional)

3.- Formulacions ideals i fortes.

Generalment hi ha més d'una forma de definir la regió factible de (PLE). Quan expressem un problema de (PLE) a través d'una sèrie de constriccions, estem definint la regió factible del nostre problema, però no és l'única forma de fer-ho. Si formulem un altre problema amb unes altres constriccions tindrem un políedre diferent → regió factible de la RL diferent, però el conjunt de solucions enteres incloses en aquell políedre seran les mateixes.

Què és el que caracteritza aquestes constriccions que ens permeten formular un problema de maneres alternatives? Totes aquestes formulacions estan construïdes a partir del que anomenem desigualtats vàlides.

Definició. Desigualtat vàlida

Una desigualtat vàlida és una constricció de desigualtat que és satisfeta per tots els punts factibles d'un problema de PLE.

↳ $a'_j x \leq b_j$ desigualtat vàlida si $a'_j x \leq b_j \ \forall x \in K_{PE}$

⊗ (PE1) i (PE2) són formulacions vàlides si $K_{PE} = K_{PE1} = K_{PE2}$ $(x^*_{PE} = x^*_{PE1} = x^*_{PE2})$

↳ Quina formulació és millor? la formulació tal que la relaxació lineal és més petita és més forta (→ millor).

Definició. Fortalesa de formulacions vàlides

Direm que la formulació PE1 és més forta que PE2 si $K_{RL1} \subset K_{RL2}$ ⇒

⇒ la relaxació lineal RL1 proporciona una fita millor (més forta) de z^*_{PE}.

⚠ si no podem demostrar que una relaxació lineal està inclosa en l'altra aleshores no les podem comparar

Definició. Formulació ideal

La formulació més forta possible de (PE) s'anomena formulació ideal (PEI)

↳ (PEI) formulació vàlida tq. $K_{RLI} \subseteq K_{RLj}$ $\forall$ formulació vàlida (PE_j) de (PE)

K_{RLI} coincideix amb l'embolcall convex de K_{PE}

$$CH(K_{PE}) \quad K_{RLI} = CH(K_{PE}) = \left\{ \sum_{i=1}^{K} \lambda_i x^i \;\middle|\; \sum_{i=1}^{K} \lambda_i = 1,\ \lambda_i \geq 0,\ x^i \in K_{PE} \right\}$$

⊗ El problema (PE) $\min\{c'x \mid x \in K_{PE}\}$ equival a (P) $\min\{c'x \mid x \in CH(K_{PE})\}$

4. Algorismes de PLE

Tenim tres algorismes per a la resolució de problemes de progr. lineal entera.

- Branch and Bound (ramifica i poda): es basa en la identificació de solucions enteres del problema de x^*_{PE} després de visitar un conjunt reduït de solucions PE usant fites z^*_{RL}
- Cutting planes (plans de tall): afegeix constriccions addicionals (talls) fins aconseguir la formulació ideal de (PE)
- Branch and cut (ramifica i talla): una combinació dels anteriors

ALGORISME DE BRANCH AND BOUND

Exemple.

$$(PE)\begin{cases} \min z_{PE} = -3x_1 - 5x_2 \\ \text{s.a.:} \quad x_1 + x_2 \geq 2 \\ \quad 7x_1 + 10x_2 \leq 35 \\ \quad x \geq 0, \text{ entera} \end{cases}$$

incumbent: representa el millor valor conegut de z_{PE1} sobre una solució factible entera

f. objectiu de PE1

Inicialització: $L = \{PE1\}$; $z^*_{PEI} = -\infty \leq z^*_{PEI} \leq z^* = +\infty$

Iteració 1)

- Seleccionem un problema $PE_j \in L$ → Agafem PE1
- Resolem la relaxació lineal RL_j: $x^*_{RL_j}$, $z^*_{PEI} \leftarrow z^*_{RL_j}$

↳ passem el problema a forma estàndard

treiem la condició d'integritat de les variables

$$(PL)\begin{cases} \min z = -3x_1 - 5x_2 \\ \text{s.a.:} \quad x_1 + x_2 - x_3 = 2 \\ \quad 7x_1 + 10x_2 + x_4 = 35 \\ \quad x \geq 0 \end{cases} \xrightarrow{\text{Fase I}} (PL)_I \begin{cases} \min z_I = x_5 \\ \text{s.a.:} \\ \quad x_1 + x_2 - x_3 + x_5 = 2 \\ \quad 7x_1 + 10x_2 + x_4 = 35 \\ \quad x \geq 0 \end{cases}$$

Resolem $(PL)_I$: $\mathcal{B} = \{5, 4\}$ $\mathcal{N} = \{1, 2, 3\}$

$B = Id = B^{-1}$ $\quad x_B = \begin{pmatrix} 2 \\ 35 \end{pmatrix} \begin{matrix} \leftarrow x_5 \\ \leftarrow x_4 \end{matrix}$ $\quad z_I = x_5 = 2$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [0 \ 0 \ 0] - [1 \ 0] \, Id \begin{bmatrix} 1 & 1 & -1 \\ 7 & 10 & 0 \end{bmatrix} =$

$= [0 \ 0 \ 0] - [1 \ 1 \ -1] = [\underset{r_1}{-1} \ \ \underset{r_2}{-1} \ \ \underset{r_3}{1}]$ $\quad q = 1$ → Entra x_1

$d_B = -B^{-1} \cdot A_q = -Id \begin{pmatrix} 1 \\ 7 \end{pmatrix} = \begin{pmatrix} -1 \\ -7 \end{pmatrix}$

$\theta^* = \min \left\{ \frac{-2}{-1}, \frac{-35}{-7} \right\} = \min \{2, 5\} = \boxed{2}$ → surt x_5

Actualització: $\mathcal{B} = \{1, 4\}$ $\mathcal{N} = \{2, 3, 5\}$

$x_B = \begin{pmatrix} 2 \\ 35 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ -7 \end{pmatrix} = \begin{pmatrix} 0 \\ 21 \end{pmatrix}$ $\quad x_1 = \theta^* = 2$

$z_I = 2 + 2(-1) = 0$ → PL factible

$B = \begin{pmatrix} 1 & 0 \\ 7 & 1 \end{pmatrix}$ $\quad B^{-1} = \begin{pmatrix} 1 & 0 \\ -7 & 1 \end{pmatrix}$ $\quad x_B = \begin{pmatrix} 2 \\ 21 \end{pmatrix}$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [0 \ 0 \ 1] - \underbrace{\cancel{[0 \ 0] B^{-1} A_N}}_{0} = [0 \ 0 \ 1] \geq [0] \Rightarrow$

⇒ Hem trobat l'òptim.

$\mathcal{B} = \{1, 4\}$ $\mathcal{N} = \{2, 3\}$ $\quad x_B = \begin{pmatrix} 2 \\ 21 \end{pmatrix}$ $\quad z = -3x_1 - \cancel{5x_2} = -6.$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [-5\ 0] - [-3\ 0]\begin{bmatrix} 1 & 0 \\ -7 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 10 & 0 \end{bmatrix} = [\underset{\uparrow\, r_2}{-2}, \underset{\uparrow\, r_3}{-3}]$

$q = 3 \rightarrow$ entra x_3

$d_B = -B^{-1} \cdot A_q = -\begin{pmatrix} 1 & 0 \\ -7 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -7 \end{pmatrix} \begin{matrix} \leftarrow d_1 \\ \leftarrow d_4 \end{matrix}$

$\theta^* = \min\left\{ \frac{-21}{-7} \right\} = 3 \rightarrow$ surt x_4

Actualització: $\mathcal{B} = \{1, 3\}$ $\mathcal{N} = \{2, 4\}$

$x_B = \begin{pmatrix} 2 \\ 21 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -7 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$ $x_3 = \theta^* = 3$

$z = -6 + (-3)3 = -15$

$\mathcal{B} = \{1, 3\}$ $x_B = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$ $B = \begin{pmatrix} 1 & -1 \\ 7 & 0 \end{pmatrix}$ $B^{-1} = \begin{pmatrix} 0 & 1/7 \\ -1 & 1/7 \end{pmatrix}$ $z = -15$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [-5\ 0] - [-3\ 0]\begin{bmatrix} 0 & 1/7 \\ -1 & 1/7 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 10 & 1 \end{bmatrix} =$

$= [-5\ 0] - [-\frac{30}{7}\ \ -\frac{3}{7}] = [\underset{\uparrow\, r_2}{-5/7}\ \ \underset{\uparrow\, r_4}{3/7}] \rightarrow q = 2$ Entra x_2

$d_B = -B^{-1} \cdot A_q = -\begin{pmatrix} 0 & 1/7 \\ -1 & 1/7 \end{pmatrix}\begin{pmatrix} 1 \\ 10 \end{pmatrix} = \begin{pmatrix} -10/7 \\ -3/7 \end{pmatrix} \begin{matrix} \leftarrow d_1 \\ \leftarrow d_3 \end{matrix}$

$\theta^* = \min\left\{ \frac{-5}{-10/7}, \frac{-3}{-3/7} \right\} = \min\left\{ \frac{35}{10}, 7 \right\} = 3'5 \rightarrow$ surt x_1

Actualització $\mathcal{B} = \{2, 3\}$ $\mathcal{N} = \{1, 4\}$

$x_B = \begin{pmatrix} 5 \\ 3 \end{pmatrix} + \frac{7}{2}\begin{pmatrix} -10/7 \\ -3/7 \end{pmatrix} = \begin{pmatrix} 0 \\ 3/2 \end{pmatrix}$ $x_2 = \theta^* = 7/2$

$z = -15 + \frac{\cancel{7}}{2}\left(\frac{-5}{\cancel{7}}\right) = -\frac{35}{2} = -17'5.$

$\mathcal{B} = \{2, 3\}$ $x_B = \begin{pmatrix} 7/2 \\ 3/2 \end{pmatrix}$ $B = \begin{pmatrix} 1 & -1 \\ 10 & 0 \end{pmatrix}$ $B^{-1} = \begin{pmatrix} 0 & 1/10 \\ -1 & 1/10 \end{pmatrix}$ $z = -17'5$

$r' = c'_N - c'_B \cdot B^{-1} \cdot A_N = [-3, 0] - [-5, 0]\begin{bmatrix} 0 & 1/10 \\ -1 & 1/10 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix} =$

$= [-3\ 0] - [-\frac{7}{2}\ \ -\frac{1}{2}] = [\frac{1}{2}\ \ \frac{1}{2}] \geq [0] \rightarrow$ hem trobat l'òptim

Un cop tenim l'òptim de RL_1 prenem una sèrie de decisions:

- Si RL és il·limitat $\Rightarrow$ PE1 il·limitat o bé infactible
- Si RL és infactible $\Rightarrow$ PE1 infactible
- Si RL té solució òptima entera $\Rightarrow$ PE1 té solució òptima

↳ Si no es dóna cap dels casos anteriors, farem una partició de la regió factible.

• separació de PE_j

Tenim $z^*_{RL1} = -17'5 \Rightarrow \lceil -17'5 \rceil = -17 = \underline{z}^*_{PEI}$

podem fer aquest arrodoniment si tots els coeficients de la f.obj són enters.

el valor de la funció objectiu de la relaxació lineal és una fita inferior del valor òptim del nostre problema

funció objectiu:

$$z_{PEI} = -3\underset{\substack{\downarrow \\ 0}}{x_1} - 5\underset{\substack{\downarrow \\ 3'5 \notin \mathbb{Z}}}{x_2}$$

$x_2^* = \frac{7}{2}$ → fem una partició de la regió factible

$$K_{PE_j} = K_{PE(k+1)} \cup K_{PE(k+2)} \quad , \quad K_{PE(k+1)} \cap K_{PE(k+2)} = \emptyset$$

✲ Ara $L = \{PE1\}$ ara hem separat PE1 en 2 problemes, per tant $L = \{PE2, PE3\}$

↳ l'òptim de PE1 és el mínim dels òptims dels dos nous problemes:

$$z^*_{PE_j} = \min\{z^*_{PE(k+1)}, z^*_{PE(k+2)}\}$$

si no tenen solució òptima entera, es separen en 2 nous problemes i així fins a trobar l'òptim.

seguim amb l'exemple:

$$x_2^* = \frac{7}{2} \rightarrow \begin{cases} (PE1) \wedge x_2 \leq \lfloor 7/2 \rfloor = 3 \rightarrow (PE2) \\ (PE1) \wedge x_2 \geq \lceil 7/2 \rceil = 4 \rightarrow (PE3) \end{cases}$$

és PE1 amb la condició que $x_2 \leq 3$ ← una constricció més

$L = \{PE2, PE3\}$

$$\underline{z}^* = -17 \leq z^*_{PEI} \leq \bar{z}^* = +\infty$$

Iteració 2)

• seleccionem PE2

• Fem la resolució de RL2

OK, perquè coefs. f.obj enters, sinó NO podríem fer l'arrodoniment.

$$(PE2) \begin{cases} \min\ z_{PE2} = -3x_1 - 5x_2 \\ \text{s.a.:} \quad x_1 + x_2 \geq 2 \\ \qquad 7x_1 + 10x_2 \leq 35 \\ \qquad x_2 \leq 3 \\ \qquad x \geq 0 \ , \ x \text{ entera} \end{cases} \quad \cdots \Rightarrow z^*_{RL2} = \lceil -17'14 \rceil = -17 \Rightarrow$$

$$\Rightarrow \underline{z}^*_{PE2} = -17$$

$x^*_{RL2} = [5/7 \quad 3] \rightarrow x_1^* \notin \mathbb{Z} \rightarrow$ fem una nova separació

$$x_1^* = \frac{5}{7} \rightarrow \begin{cases} (PE2) \wedge x_1 \leq \lfloor 5/7 \rfloor = 0 \rightarrow (PE4) \\ (PE2) \wedge x_1 \geq \lceil 5/7 \rceil = 1 \rightarrow (PE5) \end{cases}$$

$L = \{PE3, PE4, PE5\}$

Iteració 3)

- seleccionem PE4
- Resolem RL4

$$(PE4)\begin{cases} \min z_{PE4} = -3x_1 - 5x_2 \\ s.a.: \; x_1 + x_2 \geq 2 \\ 7x_1 + 10x_2 \leq 35 \\ x_2 \leq 3 \\ x_1 \leq 0 \\ x \geq 0,\; x \text{ entera} \end{cases} \quad \dots \Rightarrow z^*_{RL4} = -15$$

la solució de la relaxació lineal és entera.

$x^*_{RL4} = [0 \;\; 3] \in \mathbb{Z} \rightarrow$

→ tenim solució òptima de PE4.

- Eliminació del problema.

$L = \{PE3, PE5\}$

$z^*_{PE4} < \underset{\text{incumbent}}{z^*} \Rightarrow$ actualitzem la incumbent : $z^* = z^*_{PE4} = -15$

↳ perquè el valor de la funció objectiu és millor que el que tenia.

Per tant $\underline{z}^*_{PE1} = -17 \leq z^*_{PE1} \leq z^* = -15$, $x^* = [0 \;\; 3]$

Iteració 4)

- seleccionem PE5
- Resolem RL5

$$(PE5)\begin{cases} \min z_{PE5} = -3x_1 - 5x_2 \\ s.a.: \; x_1 + x_2 \geq 2 \\ 7x_1 + 10x_2 \leq 35 \\ x_2 \leq 3 \\ x_1 \geq 1 \\ x \geq 0;\; x \text{ entera} \end{cases} \quad \dots \Rightarrow z^*_{RL5} = -17$$

$x^*_{RL5} = [1 \;\; 2'8]$

↓

$x_2^* \notin \mathbb{Z}$

- separació:

$$x_2^* = 2'8 \rightarrow \begin{cases} (PE5) \wedge x_2 \leq \lfloor 2'8 \rfloor = 2 \rightarrow (PE6) \\ (PE5) \wedge x_2 \geq \lceil 2'8 \rceil = 3 \rightarrow (PE7) \end{cases} \quad L = \{PE3, PE6, PE7\}$$

Iteració 5)

- seleccionem PE6
- Resolem RL6 → $z^*_{RL6} = -16'43 \Rightarrow \lceil -16'43 \rceil = -16 \rightarrow \underline{z}^*_{PE6} = -16$

$x^*_{RL6} = [15/7 \;\; 2]$

↓

$x_1^* \notin \mathbb{Z}$

- Separació:

$$x_1^* = \frac{15}{7} \rightarrow \begin{cases} \overset{x_1 \leq 2}{(PE8)} \\ \underset{x_1 \geq 3}{(PE9)} \end{cases} \quad L = \{PE3, PE7, PE8, PE9\}$$

Iteració 6)

- Selecció PE8
- Relaxació PE8 → $z^*_{RL8} = -16 = z^*_{PE8}$

$$x^*_{RL8} = [2 \ \ 2] \in \mathbb{Z}$$

- Eliminació:

$x^*_{RL8} = x^*_{PE8} = [2 \ \ 2]$

$z^*_{PE8} < z^*$ → actualitzem la incumbent: $x^* = x^*_{PE8} = [2 \ \ 2]$

$z^* = z^*_{PE8} = -16$

A més, tenim que $z^* = z^*_{PE6}$ ⇒ la nostra funció objectiu té el valor de la nostra fita inferior ⇒ ja no cal mirar PE9, ja no trobarem una solució millor.

$L = \{PE3, PE7\}$

Iteració 7)

- selecció PE7
- Relaxació PE7 → $K_{RL7} = K_{PE7} = \emptyset$ infactible
- Eliminació → $L = \{PE3\}$

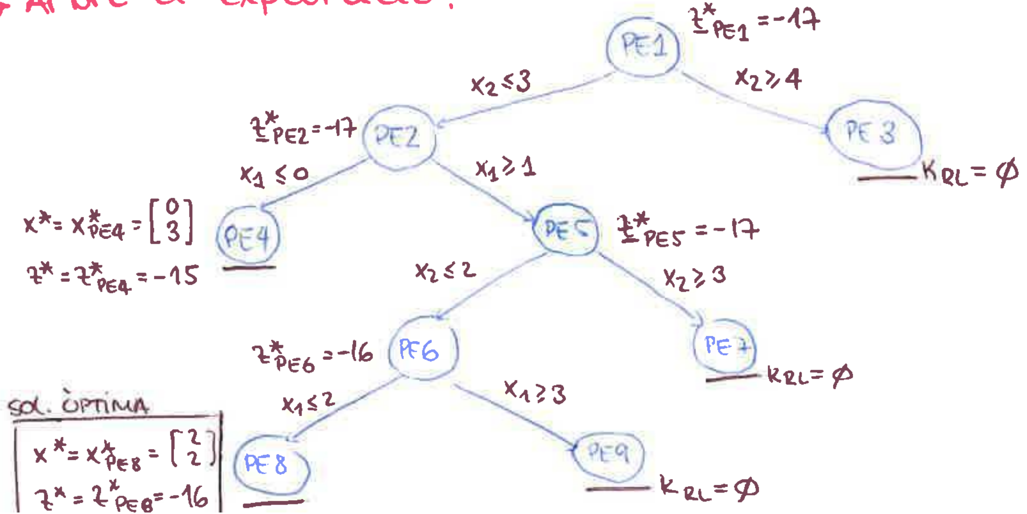
Iteració 8)

- selecció PE3
- Relaxació PE3 → $K_{RL3} = K_{PE3} = \emptyset$ infactible
- Eliminació → $L = \emptyset$ Hem acabat.

↓

$z^*_{PE1} = z^* = -16$, $x^*_{PE1} = x^* = [2, 2]$

↳ Arbre d'exploració.



Definició. Talls de Gomory.

Els talls de Gomory proporcionen un mètode sistemàtic de generació de desigualtats vàlides de problemes de PLE a partir de SBF de (RL).

La inequació $a'_j x \leq b_j$ és un tall de (PE) sobre x^*_{RL} si i només si:

i) $a'_j x \leq b_j$ és una desigualtat vàlida

ii) $a'_j x \leq b_j$ és violada per x^*_{RL}

Aplicació dels talls de Gomory.

Considerem el problema (PE) i la seva relaxació lineal:

$$(PE)\begin{cases} \min\ z = c'x \\ x \in \mathbb{Z}^n \\ s.a.:\ Ax = b \\ \quad x \geq 0 \end{cases} \rightarrow (RL)\begin{cases} \min\ z = c'x \\ x \in \mathbb{R}^n \\ s.a.:\ Ax = b \\ \quad x \geq 0 \end{cases}$$

Considerem que hem resolt (RL) amb l'algorisme del simplex obtenint la solució x^*_{RL}. Ara, tenint en compte la base associada a x^*_{RL}, reescriurem les constriccions del problema:

$$Ax = b \Rightarrow Ax = B \cdot x_B + A_N \cdot x_n = b \Rightarrow B^{-1}(Bx_B + A_N x_N) = B^{-1} \cdot b = x^*_B \Rightarrow$$

$$\Rightarrow x_B + \underbrace{(B^{-1} \cdot A_N)}_{V} x_N = x^*_B$$ → Ens fixem en l'equació i-èsima d'aquest sistema d'equacions.

$$(1)\quad x_{B(i)} + \sum_{j \in N} v_{ij} \cdot x_j = x^*_{B(i)}, \quad i \in \mathcal{B}$$ (tindrem aquesta expressió per a cadascuna de les variables bàsiques)

Agafarem aquesta relació per a generar talls.

Considerem la relació (1) associada a una component de x^*_{RL} NO entera.

Sabem que totes les solucions factibles de (PE) ($x_{PE} \in K_{PE}$) han de satisfer l'equació (1), ja que és una de les constriccions que defineixen K_{PE}.

Ara transformarem (1) en una desigualtat vàlida usant la informació que tenim sobre x_{PE} ($x_{PE} \geq 0$ i entera):

- Transformació 1: atès que $x_{PE} \geq 0$ es satisfà $\forall x_{PE} \in K_{PE}$,

$$x_{B(i)} + \sum_{j \in N} v_{ij} \cdot x_j = x^*_{B(i)} \xrightarrow{x_N \geq 0} x_{B(i)} + \sum_{j \in N} \lfloor v_{ij} \rfloor \cdot x_j \leq x^*_{B(i)}$$

- Transformació 2: atès que x_{PE} és entera es satisfà $\forall x_{PE} \in K_{PE}$,

$$x_{B(i)} + \sum_{j \in N} \lfloor v_{ij} \rfloor \cdot x_j \leq x^*_{B(i)} \xrightarrow{x \in \mathbb{Z}^n} \boxed{x_{B(i)} + \sum_{j \in N} \lfloor v_{ij} \rfloor \cdot x_j \leq \lfloor x^*_{B(i)} \rfloor}$$ ← Tall de Gomory

ALGORISME DE PLANS SECANTS DE GOMORY

Sigui (PE) amb A racional.

1.- Es resol (RL) : x^*_{RL}

2.- Si (RL) infactible o il·limitat STOP : (PE) no té solució.

3.- Si $x^*_{RL} \in \mathbb{Z}^n$, STOP : $x^*_{PE} := x^*_{RL}$

4.- Si $x^*_{RL} \notin \mathbb{Z}^n$: es selecciona una component x_i de x^*_{RL} no entera i s'afegeix a (PE) el tall de Gomory associat a x_i

5.- Anada a 1.

Exemple.

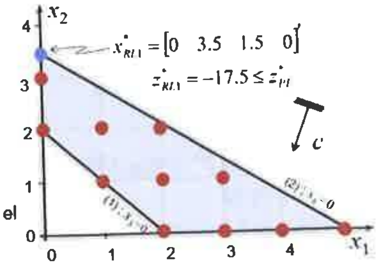
- Resoleu el següent problema amb l'algorisme de plans de tall (plans secants) secants de Gomory

$$
(PE)\begin{cases}
\min & z_{PE} = -3x_1 - 5x_2 \\
\text{s.a.:} & x_1 + x_2 \ge 2 \\
& 7x_1 + 10x_2 \le 35 \\
& x \ge 0 \text{, entera}
\end{cases}
$$

- **Alg. de plans secants de Gomory: Iteració 1**

1. Resolució de (RL1):

$$
(PE1)\begin{cases}
\min & z_{PE1} = -3x_1 - 5x_2 \\
\text{s.a.:} & x_1 + x_2 - x_3 = 2 \quad (1) \\
& 7x_1 + 10x_2 + x_4 = 35 \quad (2) \\
& x \ge 0 \text{, entera}
\end{cases}
$$



2. x^*_{RL1} no és entera: es defineix el tall de Gomory

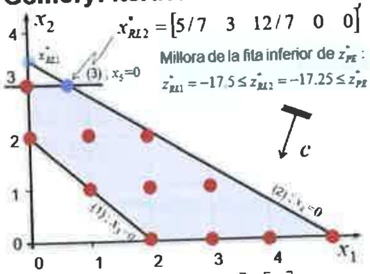
$$x_B = [x_2 \quad x_3]' = [3.5 \quad 1.5]' \text{, } V = B^{-1}A_N = \begin{bmatrix} 0 & 0.1 \\ -1 & 0.1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.1 \\ -0.3 & 0.1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$x_{B(1)} + \sum_{j\in\mathcal{N}} \lfloor v_{1j} \rfloor x_j \le \lfloor x_2^* \rfloor \rightarrow x_2 + \lfloor 0.7 \rfloor x_1 + \lfloor 0.1 \rfloor x_4 \le \lfloor 3.5 \rfloor \rightarrow \boxed{x_2 \le 3}$$

- **Alg. de plans secants de Gomory: Iteració 2**

1. Resolució de (RL2):

$$
(PE2)\begin{cases}
\min & z_{PE2} = -3x_1 - 5x_2 \\
\text{s.a.:} & x_1 + x_2 - x_3 = 2 \quad (1) \\
& 7x_1 + 10x_2 + x_4 = 35 \quad (2) \\
& x_2 + x_5 = 3 \quad \text{(3) nou tall} \\
& x \ge 0 \text{, entera}
\end{cases}
$$



2. x^*_{RL2} no és entera: es defineix el tall de Gomory

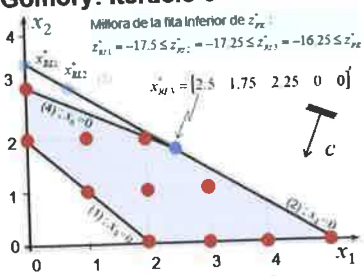
$$x_B = [x_1 \quad x_2 \quad x_3]' \text{, } V = B^{-1}A_N = \begin{bmatrix} 0 & 1/7 & -10/7 \\ 0 & 0 & 1 \\ -1 & 1/7 & -3/7 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/7 & -10/7 \\ 0 & 1 \\ 1/7 & -3/7 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$x_{B(1)} + \sum_{j\in\mathcal{N}} \lfloor v_{1j} \rfloor x_j \le \lfloor x_1^* \rfloor \rightarrow x_1 + \lfloor 1/7 \rfloor x_4 + \lfloor -10/7 \rfloor x_5 \le \lfloor 5/7 \rfloor \rightarrow x_1 - 2x_5 \le 0 \xrightarrow{(3)} \boxed{x_1 + 2x_2 \le 6}$$

- **Alg. de plans secants de Gomory: Iteració 3**

1. Resolució de (RL2):

$$
(PE3)\begin{cases}
\min & z_{PE3} = -3x_1 - 5x_2 \\
\text{s.a.} & x_1 + x_2 - x_3 = 2 \quad (1) \\
& 7x_1 + 10x_2 + x_4 = 35 \quad (2) \\
& x_2 + x_5 = 3 \quad \text{(3) redundant} \\
& x_1 + 2x_2 + x_6 = 6 \quad \text{(4) nou tall} \\
& x \ge 0 \text{, entera}
\end{cases}
$$



2. x^*_{RL2} no és entera: es defineix el tall de Gomory

$$x_B = [x_1 \quad x_2 \quad x_3]' \text{; } V = B^{-1}A_N = \begin{bmatrix} 0 & 1/2 & -5/2 \\ 0 & -1/4 & 7/4 \\ -1 & 1/4 & -3/4 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -5/2 \\ -1/4 & 7/4 \\ 1/4 & -3/4 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$x_{B(2)} + \sum_{j\in\mathcal{N}} \lfloor v_{2j} \rfloor x_j \le \lfloor x_2^* \rfloor \rightarrow x_2 + \lfloor -0.25 \rfloor x_4 + \lfloor 1.75 \rfloor x_6 \le \lfloor 1.75 \rfloor \rightarrow \boxed{x_2 - x_4 + x_6 \le 1}$$

Per tal de poder continuar resolent el problema (RL) gràficament, expressem el darrer tall de Gomory en termes de les variables x_1 i x_2 usant les constriccions (2) i (4) de (PE3):

$$
(PE3)\begin{cases}
\min & z_{PE3} = -3x_1 - 5x_2 \\
\text{s.a.:} & x_1 + x_2 - x_3 = 2 \quad (1) \\
& 7x_1 + 10x_2 + x_4 = 35 \quad (2) \rightarrow x_4 = 35 - 7x_1 - 10x_2 \\
& x_1 + 2x_2 + x_6 = 6 \quad (4) \rightarrow x_6 = 6 - x_1 - 2x_2 \\
& x \ge 0 \text{, entera}
\end{cases}
$$

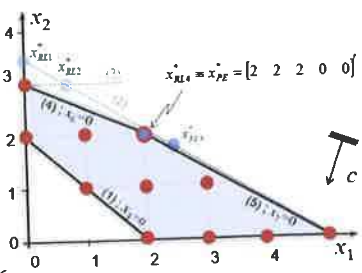
$$-x_4 + x_6 = 6x_1 + 8x_2 - 29$$

$$x_2 - x_4 + x_6 \le 1 \rightarrow 6x_1 + 9x_2 \le 30 \rightarrow \boxed{2x_1 + 3x_2 \le 10}$$

- **Alg. de plans secants de Gomory: Iteració 4**

1. Resolució de (RL2):

$$
(PE4)\begin{cases}
\min & z_{PE4} = -3x_1 - 5x_2 \\
\text{s.a.} & x_1 + x_2 - x_3 = 2 \quad (1) \\
& 7x_1 + 10x_2 + x_4 = 35 \quad \text{(2) redundant} \\
& x_2 + x_5 = 3 \quad \text{(3) redundant} \\
& x_1 + 2x_2 + x_6 = 6 \quad (4) \\
& 2x_1 + 3x_2 + x_7 = 10 \quad \text{(5) nou tall} \\
& x \ge 0 \text{, entera}
\end{cases}
$$



2. x^*_{RL2} entera: solució òptima

$$x^*_{PE} \equiv x^*_{RL4} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad , \quad z^*_{PE} \equiv z^*_{RL4} = -16$$

Comentaris:

- Les formulacions de (PE) a cada iteració son cada vegada més fortes: $z^*_{RL1} = -17.5 \le z^*_{RL2} = -17.25 \le z^*_{RL3} = -16.25 \le z^*_{RL4} = -16 \equiv z^*_{PE}$
- En aquest exemple, (PE4) és la formulació ideal: (PE4) = (PEI) = CH(K_{PE})

ALGORISME DEL BRANCH AND CUT

Algorisme genèric de Branch and Cut (B&C)

Inicialització: $L = \{PE1\}$, A racional; $z^* = +\infty$ (incumbent, $z^*_{PE1} < z^*$); $\underline{z}^*_{PE1} = -\infty$.

Mentre $L \neq \emptyset$ **fer**

Es selecciona un problema $PEj \in L$ (Selecció)

Es resol la relaxació lineal RLj **d'una formulació reforçada de PEj** : x^*_{RLj} ; $\underline{z}^*_{PEj} \leftarrow z^*_{RLj}$ (Relaxació)

Si RLj il·limitat ó $K_{RLj} = \emptyset$ ó $\underline{z}^*_{PEj} \geq z^*$ ó $x^*_{RLj} \in K_{PEj}$ **fer** (Eliminació)

$L \leftarrow L \backslash \{PEj\}$

Si $x^*_{RLj} \in K_{PEj}$ **fer**

Si $z^*_{RLj} \leq z^*$: $x^* \leftarrow x^*_{RLj}$, $z^* \leftarrow z^*_{RLj}$

Si $z^* \leq \underline{z}^*_{PEi}$ per a algun PEi : $L \leftarrow L \backslash \{PEl\}\ \forall (PEl)$ descendent de PEi

Fi Si

Altrament (Separació de PEj): sigui $k := |L|$. Es selecciona $x^*_{RLj_i} \notin \mathbb{Z}$:

Es separa PEj en els subproblemes $PE(k+1)$, $PE(k+2)$:

$(PE(k+1))\ \min\{c'x \mid x \in K_{PE(k+1)}\}$, $\quad K_{PE(k+1)} := \{x \in K_{PEj} \mid x_i \leq \lfloor x^*_{RLj_i} \rfloor\}$

$(PE(k+2))\ \min\{c'x \mid x \in K_{PE(k+2)}\}$, $\quad K_{PE(k+2)} := \{x \in K_{PEj} \mid x_i \geq \lceil x^*_{RLj_i} \rceil\}$

$K_{PEj} = K_{PE(k+1)} \cup K_{PE(k+2)}$; $K_{PE(k+1)} \cap K_{PE(k+2)} = \emptyset$; $\underline{z}^*_{PEj} = \min\{\underline{z}^*_{PE(k+1)}, \underline{z}^*_{PE(k+2)}\}$

Es substitueix PEj pels seus descendents : $L \leftarrow L \backslash \{PEj\} \cup \{PE(k+1)\} \cup \{PE(k+2)\}$

Fi Si

Fi Mentre

Si $z^* < +\infty$: $x^*_{PE1} \equiv x^*$.
Altrament Si $RL1$ il·limitat : $PE1$ il·lim. o infac.
Altrament : $PE1$ infactible.
Fi Si

$$(RL1)\begin{cases} \text{Sol. opt.} \Rightarrow (PE1)\begin{cases}\text{Sol. opt.} & \Rightarrow \exists (PEj) \text{ sol. opt. } (z^* < +\infty). \\ \text{Infac.} & \Rightarrow \text{Tot } (PEj) \text{ infac. } (z^* = +\infty).\end{cases} \\ \text{Infac.} \Rightarrow (PE1) \text{ infactible.} \\ \text{Il·lim.} \Rightarrow (PE1) \text{ infactible o il·limitat.} \end{cases}$$

TEMA 5 : Optimització no lineal sense restriccions

1.- Conceptes bàsics

Definició. Problema d'optimització no lineal (ONL)

Considerem la funció $f(x)$ i el conjunt Ω, on $f: \mathbb{R}^n \to \mathbb{R}$ és la funció objectiu (suposem "suau", $f \in \mathcal{C}^2$) i $\Omega \subseteq \mathbb{R}^n$ és el conjunt factible.

Per tant en el nostre problema voldrem $\left.\begin{array}{l} \min f(x) \\ \text{s.a.: } x \in \Omega \end{array}\right\}$

- si $\Omega = \mathbb{R}^n$ tenim un problema sense restriccions $\rightarrow$ $\left.\begin{array}{l} \min f(x) \\ x \in \mathbb{R}^n \end{array}\right\}$
- si $\Omega \subset \mathbb{R}^n \Rightarrow \Omega = \{x \mid h(x) = 0,\ g(x) \leq 0\}$

on $h: \mathbb{R}^n \to \mathbb{R}^m$ $(h_i: \mathbb{R}^n \to \mathbb{R})$ $i = 1, \dots, m$, $g: \mathbb{R}^n \to \mathbb{R}^p$ $(g_j: \mathbb{R}^n \to \mathbb{R})$ $j = 1, \dots, p$ suposem $h, g \in \mathcal{C}^2$

tenim un problema amb restriccions d'igualtat i/o desigualtat $\rightarrow$ $\left.\begin{array}{l} \min f(x) \\ \text{s.a.: } h(x) = 0 \\ \quad\quad g(x) \leq 0 \end{array}\right\}$

※ Considerarem problemes de minimització. Notem que

$$\max f(x),\ x \in \Omega \ \equiv\ -\min -f(x),\ x \in \Omega$$

⚠ **Notació.**

$f(x^*) = \min\limits_{x \in \Omega} f(x)$ $\rightarrow$ valor de la funció amb x òptima

$x^* = \arg\min\limits_{x \in \Omega} f(x)$ $\rightarrow$ valor de la x òptima

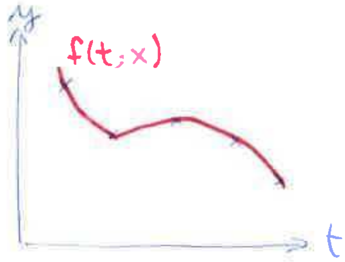
※ Els mètodes de resolució també es classifiquen en mètodes per a problemes d'ONL amb i sense restriccions. És una de les principals diferències amb PL's, ja que els PL's sense restriccions són il·limitats.

Exemples. Modelització.

1) Ajust de corbes per mínims quadrats no lineals.

Tenim m punts de dades (t_i, y_i) on $t_i \in \mathbb{R}^k$, $y_i \in \mathbb{R}^l$

Per exemple, si $k = l = 1$ podríem tenir



Volem ajustar una funció $f(t;x): \mathbb{R}^k \to \mathbb{R}^l$, depenent d'un vector de paràmetres $x \in \mathbb{R}^n$

(donada una t, hem de predir y en funció d'unes x)

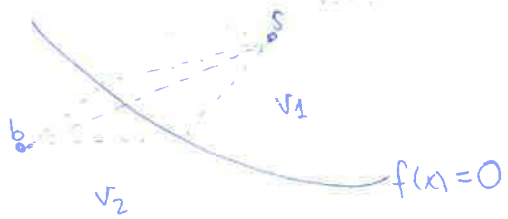
Com trobem la millor corba o model?

Minimitzem els errors: $r_i = f(t_i; x) - y_i \quad i = 1, \dots, m$

↑ error que estic cometent en el punt i

Formulem: $$\min_{x \in \mathbb{R}^n} \sum_{i=1}^{n} \|r_i\|^2 = \min_{x \in \mathbb{R}^n} \sum_{i=1}^{n} (f(t_i; x) - y_i)^T (f(t_i; x) - y_i)$$

2) Problemes de distàncies. La socorrista i el banyista



La socorrista està al punt $s = (s_1, s_2)$, el banyista està al punt $b = (b_1, b_2)$ $\Big\}$ $s, b \in \mathbb{R}^2$

La frontera entre la platja i el mar són els punts x tq $f(x) = 0$.

La socorrista corre per la sorra i neda a unes velocitats v_1 m/s i v_2 m/s respectivament.

Volem saber quin recorregut ha de fer la socorrista per arribar al banyista en el temps mínim.

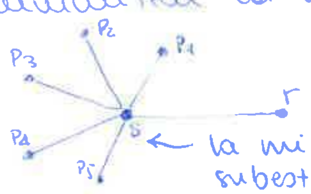
↳ Cal trobar $x^* \in \mathbb{R}^2$ que proporciona el temps mínim i verifica $f(x^*) = 0$.

$$\left.\begin{aligned} \min_{x \in \mathbb{R}^2} \; & t_1(x) + t_2(x) \\ \text{s.a.: } & f(x) = 0 \end{aligned}\right\} \quad \text{on} \quad \begin{aligned} t_1(x) &= \frac{d(x,s)}{v_1} = \frac{\sqrt{(s_1 - x_1)^2 + (s_2 - x_2)^2}}{v_1} \\ t_2(x) &= \frac{d(x,b)}{v_2} = \frac{\sqrt{(b_1 - x_1)^2 + (b_2 - x_2)^2}}{v_2} \end{aligned}$$

3) Problemes de distàncies. L'oleoducte.

Es vol construir un oleoducte per comunicar n pous de petroli de coordenades $p_i = (x_i, y_i) \in \mathbb{R}^2$, amb una refineria de coordenades $r = (x_r, y_r)$, $r \in \mathbb{R}^2$

Volem minimitzar la longitud de l'oleoducte.



← la millor solució passa per unir els diferents pous en una subestació $s = (x_s, y_s) \in \mathbb{R}^2$.

Per a trobar situació de s formulem el problema sense restriccions:

$$\min_{s\in\mathbb{R}^2} \sum_{i=1}^{n} \|s-p_i\|_2 + \|r-s\|_2 = \min_{s\in\mathbb{R}^n} \sum_{i=1}^{n} \sqrt{(x_s-x_i)^2+(y_s-y_i)^2} + \sqrt{(x_r-x_s)^2+(y_r-y_s)^2}$$

4) Càlcul de centres. Centre de Chebyshev

↳ volem el radi de la circumferència més gran que podem inscriure en el conjunt C.

$C \subseteq \mathbb{R}^n$ fitat

↳ si és un conjunt qualsevol és un problema difícil.

↳ si és un conjunt convex definit com $C = \{x\in\mathbb{R}^n \mid f_i(x) \leq 0 \quad i=1,\dots,m\}$

$f_1,\dots,f_m$ funcions convexes

Aleshores podem formular el problema:

Volem maximitzar el radi.

Variables: $x\in\mathbb{R}^n$ centre, $r\in\mathbb{R}$ radi $\Big\}$ $\max_{x,r} r$ tq $r \geq 0$.

Hem de garantir que tots els punts del cercle compleixen les restriccions. $x + r\cdot u \in C$

u: vector $\in \mathbb{R}^n$ tq $\|u\| \leq 1$

$$g_i(x) = \max_{\|u\|\leq 1} f(x+r\cdot u)$$

↳ optimitzem respecte u, x i r són paràmetres

$g_i(x,r) \leq 0 \quad i=1,\dots,m$

↳ Es tracta d'un problema d'optimització binivell, és a dir, plantegem un problema d'optimització i les seves restriccions també són un problema d'optimització.

Definició. Gradient, Hessiana i Jacobiana.

Sigui $f: D\subseteq\mathbb{R}^n \to I \subseteq \mathbb{R}$, $f\in\mathcal{C}^2$. El vector gradient de f és:

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

La matriu Hessiana de f és:

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

Sigui $g: D\subseteq\mathbb{R}^n \to I\subseteq\mathbb{R}^m$, $g=(g_1,\dots,g_m)^T$, $g_i\in\mathcal{C}$. La matriu Jacobiana és:

$$\nabla g(x)^T = J(x) = \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

Exemple.

$f(x) = e^{x_1^2 + x_2}$

$$\nabla f(x) = \begin{bmatrix} 2x_1 e^{x_1^2+x_2} \\ e^{x_1^2+x_2} \end{bmatrix} \begin{array}{l} \leftarrow \frac{\partial f}{\partial x_1} \\ \leftarrow \frac{\partial f}{\partial x_2} \end{array}$$

$$\nabla^2 f(x) = \begin{bmatrix} 2e^{x_1^2+x_2} + 4x_1^2 e^{x_1^2+x_2} & 2x_1 e^{x_1^2+x_2} \\ 2x_1 e^{x_1^2+x_2} & e^{x_1^2+x_2} \end{bmatrix}$$

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ 3x_1 + 4x_2 - 5 \end{bmatrix}$$

$$\nabla g(x)^T = J(x) = \begin{bmatrix} \nabla g_1(x)^T \\ \nabla g_2(x)^T \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \\ 3 & 4 \end{bmatrix}$$

Definició. Teorema de Taylor

Sigui $f: D \subseteq \mathbb{R} \to I \subseteq \mathbb{R}$ tq $f \in \mathcal{C}^k$. Sigui $\alpha \in \mathbb{R}$. Aleshores:

$$f(x+\alpha) = f(x) + f'(x)\cdot\alpha + \frac{f''(x)}{2!}\alpha^2 + \dots + \frac{f^{(k)}(x)}{k!}\alpha^k + \underbrace{R_k(\alpha)}_{\text{residu de Taylor}}, \quad R_k(\alpha) = o(\alpha^k)$$

La versió sense residu:

$$f(x+\alpha) = f(x) + f'(x)\alpha + \frac{f''(x)}{2!}\alpha^2 + \dots + \frac{f^{(k)}(z)}{k!}\alpha^k, \quad z = x + \theta\alpha, \ 0 \le \theta \le 1$$

Definició. Polinomi de Taylor

Sigui $f: D \subseteq \mathbb{R}^n \to I \subseteq \mathbb{R}$, $f \in \mathcal{C}^2$, $d \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$

- Polinomi d'ordre 1:

  $f(x+\alpha d) = f(x) + \alpha \nabla f(x)^T d + R_1(\alpha)$, $R_1(\alpha) = o(\alpha)$ → amb residu

  $f(x+\alpha d) = f(x) + \alpha \nabla f(z)^T d$, $z = x + \theta\alpha d$, $0 \le \theta \le 1$ → sense residu

- Polinomi d'ordre 2:

  $f(x+\alpha d) = f(x) + \alpha \nabla f(x)^T d + \frac{1}{2}\alpha^2 d^T \nabla^2 f(x) d + R_2(\alpha)$, $R_2(\alpha) = o(\alpha^2)$

  $f(x+\alpha d) = f(x) + \alpha \nabla f(x)^T d + \frac{1}{2}\alpha^2 d^T \nabla^2 f(z) d$, $z = x + \theta\alpha d$, $0 \le \theta \le 1$

⚠ $f(t) = o(g(t))$ si $\lim_{t\to 0} \frac{f(t)}{g(t)} = 0$ ($f(t)$ és infinitèssim d'ordre superior)

Teorema. Teorema de Weierstrass (existència d'òptim)

Si $f: \Omega \to \mathbb{R}$ és contínua i $\Omega \subseteq \mathbb{R}^n$ és un conjunt compacte (tancat i fitat) aleshores $\exists x^* \in \Omega$ tq $f(x^*) \le f(x)$ $\forall x \in \Omega$ (és a dir, $\min_{x\in\Omega} f(x)$ té solució)

Definició. Mínim local i mínim global

Donat el problema d'optimització $\min f(x)$
$\text{s.a.: } x \in \Omega$

diem que x^* és el mínim (òptim) local si $x^* \in \Omega$ i hi ha un entorn B de x^* (conjunt obert que conté x^*) tal que $f(x^*) \leq f(x) \; \forall x \in B \cap \Omega$ (es diu estricte si $f(x^*) < f(x)$).

diem que x^* és el mínim (òptim) global si $x^* \in \Omega$ i $f(x^*) \leq f(x) \; \forall x \in \Omega$

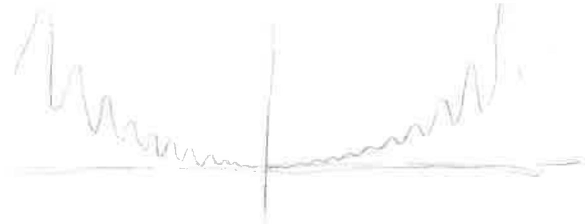
Definició. Mínim local aïllat

x^* és mínim local aïllat si hi ha un entorn B de x^* on x^* és l'únic mínim

⚠ **Els mínims locals estrictes no sempre són aïllats.** (min local aïllat $\Rightarrow$ min local estricte; $\not\Leftarrow$)

Exemple.

$$f(x) = \begin{cases} x^4 \cdot \cos\left(\frac{1}{x}\right) + 2x^4 & x \neq 0 \\ 0 & x = 0 \end{cases}$$



Té un mínim local estricte a $x^* = 0$, però no és mínim aïllat: hi ha molts mínims locals estrictes tan a prop com vulguem de $x^* = 0$.

2.- Convexitat

Sempre desitgem obtenir l'òptim global, però a la pràctica això no és possible. Els mètodes numèrics de ONL es basen en condicions d'optimalitat locals. Hi ha un tipus de problemes per als quals podem garantir òptims globals, els problemes convexos.

Definició. Conjunt convex.

$\Omega \subseteq \mathbb{R}^n$ és un conjunt convex si $\forall x, y \in \Omega \quad \alpha x + (1-\alpha) y \in \Omega, \; 0 \leq \alpha \leq 1.$

(Gràficament vol dir que el segment $\overline{xy} \in \Omega$)

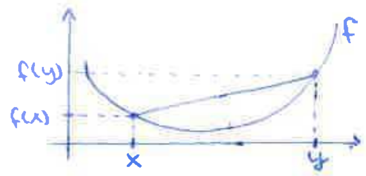
convex — no convex

Definició. Funció convexa.

f és convexa en $\Omega \subseteq \mathbb{R}^n$ si $\forall x, y \in \Omega \quad f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

$\alpha \in [0, 1]$

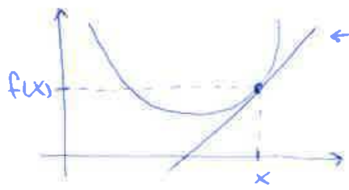
Interpretació geomètrica:



Proposició. Caracterització de funcions convexes usant ∇f.

$f \in \mathcal{C}^1$ és convexa en $\Omega \subseteq \mathbb{R}^n \iff f(y) \geq f(x) + \nabla f(x)^T(y-x) \quad \forall x,y \in \Omega$

Interpretació geomètrica:



La convexitat és estricta $\iff$ la desigualtat és $>$ $\forall x,y$ tq $x \neq y$

Demostració.

Volem demostrar que (1) i (2) són equivalents on

(1): $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \quad \forall \alpha \in [0,1]$

(2): $f(y) \geq f(x) + \nabla f(x)^T(y-x)$

(1) $\Rightarrow$ (2) | $f(x+\alpha(y-x)) \leq f(x) + \alpha(f(y)-f(x))$

$$\frac{f(x+\alpha(y-x)) - f(x)}{\alpha} \leq f(y) - f(x)$$

Si fem tendir α a 0 tenim la definició de derivada direccional

per tant $\nabla f(x)^T(y-x) \leq f(y) - f(x) \Rightarrow$

$\Rightarrow f(y) \geq f(x) + \nabla f(x)^T(y-x)$ OK

(2) $\Rightarrow$ (1) | considerem $x_\alpha = y + \alpha(x-y) \in \Omega$ ($x_\alpha = \alpha x + (1-\alpha)y$)

$(\alpha \in [0,1])$

(x_α, x) $\quad [\alpha][f(x) \geq f(x_\alpha) + \nabla f(x_\alpha)^T(x-x_\alpha)]$

(x_α, y) $\quad [1-\alpha][f(y) \geq f(x_\alpha) + \nabla f(x_\alpha)^T(y-x_\alpha)]$

Fem la suma $(x_\alpha, x) + (x_\alpha, y)$

$$(\alpha) f(x) + (1-\alpha) f(y) \geq \underbrace{\alpha f(x_\alpha) + (1-\alpha) f(x_\alpha)}_{f(x_\alpha)} + \nabla f(x_\alpha)^T(\underbrace{\alpha(x-x_\alpha)+(1-\alpha)(y-x_\alpha)}_{0})$$

$$\alpha(x-x_\alpha) + (1-\alpha)(y-x_\alpha) =$$

$$= \alpha x + (1-\alpha)y - x_\alpha = x_\alpha - x_\alpha = 0.$$

Per tant, $\alpha f(x) + (1-\alpha) f(y) \geq f(x_\alpha) \Rightarrow$

$\Rightarrow \alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y)$ OK

⚠ Si tenim desigualtats estrictes (2) $\Rightarrow$ (1) es compleix però (1) $\Rightarrow$ (2) NO es compleix.

Definició. Matriu definida positiva i semidefinida positiva

Una matriu simètrica $H \in \mathbb{R}^{n\times n}$ és semidefinida positiva (definida positiva) si $\forall x \in \mathbb{R}^n$, $x \neq 0$, $x^T \cdot H x \geq 0$ $(x^T \cdot H \cdot x > 0)$

una matriu és semidefinida positiva (definida positiva) $\Leftrightarrow$

$\Leftrightarrow$ $\begin{cases} \bullet \text{ Tots els seus VAPS són } \geq 0 \quad (>0) \\ \bullet \text{ Tots els menors principals són } \geq 0 \text{ (els n menors dominants són } >0) \end{cases}$

Proposició. Caracterització de les funcions convexes usant $\nabla^2 f$

$f \in \mathcal{C}^2$ és convexa en $\Omega \subseteq \mathbb{R}^n$ convex i que conté com a mínim un punt interior $\Leftrightarrow$ $\nabla^2 f$ és semidefinida positiva en Ω. $(\nabla^2 f \succcurlyeq 0$ en $\Omega)$

si $\nabla^2 f \succ 0$ la convexitat és estricta (condició suficient, no necessària)

Demostració

- $\nabla^2 f \succcurlyeq 0 \Rightarrow f$ convexa:

tenim $x, y \in \Omega$ tq. $f(y) = f(x) + \nabla f(x)^T(y-x) + \underbrace{\frac{1}{2}(y-x)^T \nabla^2 f(z)}_{\geq 0}$ $\Rightarrow$

escrivim la sèrie de taylor de 2n ordre sense residu calculant y a partir de x.

x z y

$z = x + \theta(y-x)$
$\theta \in (0,1)$

$z \in \Omega$ (convexitat de Ω)
$\Downarrow$
$\nabla^2 f(z) \succcurlyeq 0$

$\Rightarrow f(y) \geq f(x) + \nabla f(x)^T(y-x) \Rightarrow f$ convexa ok
(per la proposició anterior)

- f convexa $\Rightarrow \nabla^2 f \succcurlyeq 0$: per reducció a l'absurd.

Suposem f convexa i $\nabla^2 f \not\succcurlyeq 0$ en Ω $\Rightarrow$

$\Rightarrow \exists x \in \Omega \mid$ per a algun d, $d^T \cdot \nabla^2 f(x) \cdot d < 0$

Per continuïtat de la Hessiana puc suposar que x és interior a Ω.

Considerem $y = x + \alpha d$, $y - x = \alpha d$ $(\alpha \geq 0)$

$(y-x)^T \nabla^2 f(x)(y-x) = \underbrace{\alpha^2}_{\geq 0} \underbrace{d^T \nabla^2 f(x) d}_{< 0} < 0$

$\forall z = x + \theta(y-x)$ $\Rightarrow$ $(y-x)^T \nabla^2 f(z)(y-x) < 0$
$\theta \in (0,1)$

Per teorema de Taylor, $f(y) = f(x) + \nabla f(x)^T(y-x) + \underbrace{\frac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)}_{< 0}$ $\Rightarrow$

$\Rightarrow f(y) < f(x) + \nabla f(x)^T(y-x)$

No es verifica la desigualtat de la proposició anterior $\Rightarrow f$ NO convexa
$\downarrow$
CONTRADICCIÓ!

Proposició. Caracterització de conjunts convexos.

Sigui $g: D \to \mathbb{R}$ (D convex) una funció convexa. Aleshores el conjunt

$S_c = \{x \mid x \in D, g(x) \leq c\}$ és convex $\forall c \in \mathbb{R}$.

Demostració: EXERCICI.

⊗ Com a conseqüència del resultat anterior, el conjunt $\Omega = \{x \mid g_j(x) \leq 0, j = 1, \dots, p\}$ és convex si g_j són funcions convexes, donat que la intersecció de conjunts convexos és un conjunt convex.

El conjunt $\Omega = \{x \mid h_i(x) = 0, i = 1, \dots, m\}$ és convex si h_i són funcions afins, és a dir, $h_i(x) = a^T x - b$

La condició de la proposició anterior és suficient per a garantir una regió factible convexa, però no necessària.

Definició. Problema convex.

Un problema d'optimització $\begin{array}{l} \min f(x) \\ \text{s.a.: } x \in \Omega \end{array}$ és convex si f és una funció convexa i Ω és un conjunt convex.

Teorema. Donat $\min f(x)$ s.a.: $x \in \Omega$ un problema convex. Aleshores:

i) Si x^* és mínim local de f en Ω $\Rightarrow$ x^* és mínim global de f en Ω.

ii) Si f és estrictament convexa $\Rightarrow$ només hi ha un únic mínim global de f en Ω.

Demostració.

i) Per reducció a l'absurd.

Suposem x^* mínim local que NO és mínim global. $\Rightarrow$

$\Rightarrow \exists z \in \Omega \mid f(z) < f(x^*)$.

Considerem $f(\alpha x^* + (1-\alpha) z) \underset{\substack{\uparrow \\ f \text{ convexa}}}{\leq} \alpha f(x^*) + (1-\alpha) \overbrace{f(z)}^{< f(x^*)} < \alpha f(x^*) + (1-\alpha) f(x^*) = f(x^*)$

Per tant hi ha punts del segment tant a prop com vulguem de x^* que són millors que x^* $\Rightarrow$ x^* NO és mínim local CONTRADICCIÓ.

ii) Per reducció a l'absurd.

Suposem $x, y \in \Omega$ tq $x \neq y$, $f(x) = f(y)$ mínims globals.

$f(\alpha x + (1-\alpha) y) < \alpha \underset{f(y)}{\underset{\parallel}{f(x)}} + (1-\alpha) \underset{f(x)}{\underset{\parallel}{f(y)}} = f(x) = f(y) \Rightarrow$

$\Rightarrow$ hi ha punts del segment millors que x i y $\Rightarrow$ CONTRADICCIÓ.

Proposició. Convexitat del conjunt de solucions.

Sigui $\min f(x)$ s.a.: $x \in \Omega$ un problema convex (f convexa i Ω convex).

Aleshores el conjunt de solucions $\Omega^* \subseteq \Omega$ és convex.

Demostració: EXERCICI

3.- Optimització no lineal sense restriccions

S'ha de motivar les condicions d'optimalitat.

Com caracteritzem x^*? $f(x^*) \leq f(x)$ $\forall x \in \mathbb{R}^n$

Exemple.

$\min f(x_1, x_2) = x_1^2 + x_2^2$

El mínim local i global estricte és $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Un altre punt x no és mínim perquè hi ha direccions $d \in \mathbb{R}^2$ tq

$f(x + \alpha d) < f(x)$ $\alpha > 0$.

↳ Per exemple, al punt $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ al llarg de la direcció $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

tenim $f(x + \alpha d) = (1-\alpha)^2 + (1-\alpha)^2 = 2 - 4\alpha + 2\alpha^2 = f(x) - 4\alpha + 2\alpha^2 < 0$ si $\alpha \in (0,2)$

Pel T. Taylor, $f(x + \alpha d) = f(x) + \alpha \nabla f(x)^T d + o(\alpha)$

veiem que tindrem una direcció d: $f(x + \alpha d) < f(x)$ si $\nabla f(x)^T d < 0$

I sempre podrem fer $\nabla f(x)^T d < 0$ escollint per exemple $d = -\nabla f(x)$.

Per tant, la única forma d'evitar $\nabla f(x)^T d < 0$ és que $\nabla f(x) = 0$. Aquesta és la condició necessària d'optimalitat de 1r ordre.

Teorema. Condicions necessàries d'optimalitat

Sigui $\min_{x \in \mathbb{R}^n} f(x)$, $f: \mathbb{R}^n \to \mathbb{R}$.

Si x^* és un mínim local de f i $f \in \mathcal{C}^2$ en un entorn obert de x^* llavors:

i) $\nabla f(x^*) = 0$ (condició de 1r ordre)

ii) $\nabla^2 f(x^*) \geq 0$ (condició de 2n ordre).

Els punts x^* tq $f(x^*) = 0$ s'anomenen punts estacionaris.

Demostració

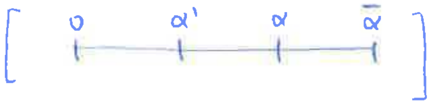
i) Per reducció a l'absurd. Suposem $\nabla f(x^*) \neq 0 \Rightarrow \exists d \in \mathbb{R}^n$ tq $\nabla f(x^*)^T \cdot d < 0$

(per exemple $d = -\nabla f(x^*)$).

Aleshores, com que $f \in \mathcal{C}^2$ (particularment $f \in \mathcal{C}^1$), per continuïtat de $\nabla f \Rightarrow \exists \bar{\alpha}$ tq. $\nabla f(x^* + \alpha d)^T \cdot d < 0 \quad \forall \alpha \in (0, \bar{\alpha})$

Prenem un $\alpha \in (0, \bar{\alpha})$. $f(x^* + \alpha d) \underset{\substack{\uparrow \\ \text{teorema} \\ \text{de taylor}}}{=} f(x^*) + \nabla f(x^* + \alpha' d)^T \cdot d$

$\alpha' \in (0, \alpha) \Rightarrow \alpha' \in (0, \bar{\alpha})$



Com que $\alpha' \in (0, \bar{\alpha})$, $\nabla f(x^* + \alpha' d)^T \cdot d < 0 \Rightarrow$

$\Rightarrow f(x^* + \alpha' d) < f(x^*) \rightarrow x^*$ <u>NO</u> és mínim $\rightarrow$ CONTRADICCIÓ

ii) Per reducció a l'absurd. Suposem que $\nabla^2 f(x^*) \not\succeq 0 \Rightarrow$

$\Rightarrow \exists d \in \mathbb{R}^n$ tq $d^T \cdot \nabla^2 f(x^*) < 0$

Com que $f \in \mathcal{C}^2$, per continuïtat de $\nabla^2 f$, $\exists \bar{\alpha}$ tq $d^T \cdot \nabla^2 f(x^* + \alpha d) \cdot d < 0 \quad \forall \alpha \in (0, \bar{\alpha})$.

Considerem $\alpha \in (0, \bar{\alpha})$. $f(x^* + \alpha d) \underset{\substack{\uparrow \\ \text{Teorema} \\ \text{de Taylor}}}{=} f(x^*) + \underbrace{\alpha \nabla f(x^*)^T d}_{(i) \rightarrow \; = 0} + \underbrace{\frac{\alpha^2}{2} \cdot d^T \nabla^2 f(x^* + \alpha d)\, d}_{< 0} \Rightarrow$

perquè $\alpha \in (0, \bar{\alpha})$

$\Rightarrow f(x^* + \alpha d) < f(x^*) \Rightarrow x^*$ <u>no</u> és mínim $\rightarrow$ CONTRADICCIÓ.

Teorema. Condicions suficients d'optimalitat.

Si $f \in \mathcal{C}^2$ en un entorn obert de x^*, $\nabla f(x^*) = 0$ i $\nabla^2 f(x^*)$ és definida positiva llavors x^* és mínim local estricte de f.

<u>Demostració</u>

$\nabla^2 f(x^*) \succ 0 \Rightarrow \forall d \in \mathbb{R}^n \quad d^T \nabla^2 f(x^*)\, d > 0$

Aleshores, com que $f \in \mathcal{C}^2$, $\nabla^2 f(x^*)$ és contínua i $\exists B_r(x^*) \mid \forall z \in B$ $d^T \nabla^2 f(z)\, d > 0$

$\forall$ punt $x^* + \alpha d$ tq $\|\alpha d\| < r$, $f(x^* + \alpha d) = f(x^*) + \alpha \nabla f(x^*)^T d + \underbrace{\frac{\alpha^2}{2} d^T \nabla^2 f(z)\, d}_{> 0}$ $\quad z \in B$

$z = x^* + \theta \alpha d$

$\theta \in (0, 1)$

$\forall d$ tq $f(x^* + \alpha d) > f(x^*) \Rightarrow x^*$ és un mínim local i estricte <u>OK</u>

Teorema. Condicions d'optimalitat en problemes convexos.

Si f és convexa i diferenciable, llavors un punt estacionari x^* ($\nabla f(x^*) = 0$) és mínim global de f.

Demostració Per reducció a l'absurd.

Suposem que x^* NO és mínim global $\Rightarrow \exists z \in \mathbb{R}^n \mid f(z) < f(x^*)$

$$0 = \underbrace{\nabla f(x^*)^T (z - x^*)}_{\substack{\text{derivada direccional} \\ \text{de } f \text{ en } x \text{ al llarg} \\ \text{de } (z-x^*)}} = \lim_{\lambda \to 0} \frac{f(x^* + \lambda(z - x^*)) - f(x^*)}{\lambda} =$$

per convexitat de f

$$= \lim_{\lambda \to 0} \frac{f(\lambda z + (1-\lambda)x^*) - f(x^*)}{\lambda} \leq \lim_{\lambda \to 0} \frac{\lambda f(z) + (1-\lambda) f(x^*) - f(x^*)}{\lambda} =$$

$$= \lim_{\lambda \to 0} \frac{\cancel{\lambda} f(z) - \cancel{\lambda} f(x^*)}{\cancel{\lambda}} = f(z) - f(x^*)$$

Per tant tenim que $0 \leq f(z) - f(x^*) \Rightarrow f(x^*) \leq f(z) \rightarrow$ CONTRADICCIÓ!

⊛ Així doncs, si $f \in \mathcal{C}^1$ és convexa només cal comprovar si el punt és estacionari; si això es verifica el punt és mínim local. I no només local, també global. A diferència de les condicions suficients no cal considerar que $\nabla^2 f(x^*)$ sigui definida positiva; en ser f convexa n'hi ha prou amb que $\nabla^2 f(x^*)$ sigui semidefinida positiva.

Exemples.

Diferents casos que es poden donar:

1) $f(x) = -e^{-(x_1^2 + x_2^2)}$

$$\nabla f(x) = \begin{bmatrix} 2x_1 e^{-(x_1^2+x_2^2)} \\ 2x_2 e^{-(x_1^2+x_2^2)} \end{bmatrix}$$

Punts estacionaris : $\nabla f(x^*) = 0$

↳ $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\nabla^2 f(x) = e^{-(x_1^2+x_2^2)} \begin{bmatrix} 2-4x_1^2 & -4x_1x_2 \\ -4x_1x_2 & 2-4x_2^2 \end{bmatrix}$$

→ Definida positiva $\Rightarrow$

$\Rightarrow x^*$ és mínim local estricte.

$$\nabla^2 f(x^*) = e^0 \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

↳ En aquest cas x^* també és mínim global, tot i que f NO és convexa.

2) $f(x) = x_1^2 - x_2^2 - 4x_1 + 6x_2 - 5$

$$\nabla f(x) = \begin{bmatrix} 2x_1 - 4 \\ -2x_2 + 6 \end{bmatrix}$$

Punts estacionaris : $\nabla f(x^*) = 0$

↳ $x^* = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

→ És indefinida, té VAPs positius i negatius.

↳ En aquest cas x^* no és mínim, és un punt de sella.

3) $f(x) = x_1^2 - x_2^4$

$\nabla f(x) = \begin{bmatrix} 2x_1 \\ -4x_2^3 \end{bmatrix}$ Punts estacionaris: $\nabla f(x^*) = 0$

↳ $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -12x_2^2 \end{bmatrix}$ → És semidefinida positiva, NO definida positiva ⇒ NO podem assegurar que sigui mínim.

↳ De fet NO és mínim, al llarg de qualsevol direcció $d = \begin{bmatrix} 0 \\ d_2 \end{bmatrix}$ f minora:

$f(x^* + \alpha d) = f([0, \alpha d]^T) = -(\alpha d_2)^4 < 0 = f(x^*)$.

Tenim un punt que satisfà les condicions necessàries i NO és mínim.

4) $f(x) = x_1^2 + x_2^4$

$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}$ Punts estacionaris: $\nabla f(x^*) = 0$

↳ $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$ → És semidefinida positiva, NO definida positiva ⇒ NO podem assegurar que sigui mínim.

↳ Però en aquest cas, tot i que no satisfà les condicions suficients, sí que és un mínim local (i estricte).

4.- Mètodes d'optimització

⊛ Per què no podem solucionar $\nabla f(x) = 0$ directament?

↳ $\nabla f(x) = 0$ és sovint un sistema d'equacions no lineal.

↳ Els punts estacionaris NO tenen perquè ser mínims. Els mètodes d'optimització estan dissenyats per a anar al mínim.

↳ Poden ser més ràpids que resoldre el sistema $\nabla f(x) = 0$.

Definició. Mètodes d'exploració lineal. (line search)

Aquests mètodes generen una seqüència de punts $\{x^k\}_0^\infty$ que (sota certes condicions) convergeix a l'òptim.

La seqüència generada és: $x^{k+1} = x^k + \underbrace{\alpha^k}_{\text{longitud de pas}} \underbrace{d^k}_{\text{direcció}}$ (primer calculem $d^k \in \mathbb{R}^n$ i després $\alpha^k \in \mathbb{R}$)

Definició. Mètodes de regió de garantia (trust region methods)

En aquests mètodes $x^{k+1} = x^k + d^k$ → resolem min $m_k(d)$, que és un model d'aproximació de $f(x)$ al punt x^k.

Es compleix que $d \in \mathbb{R}^n$, $\|d\| < \Delta_k$ → indica com de gran és la regió de garantia. És diferent en cada iteració.

↳ En aquest curs només veurem els mètodes d'exploració lineal.

Direccions d'ascens i de descens

⊛ Com sabem si x^k és solució? S'ha de complir $\nabla f(x^k) = 0$ (o bé $\|\nabla f(x^k)\| < \varepsilon$ ← tolerància prefixada)

⚠ No sempre que $\nabla f(x^k) = 0$ voldrà dir que tenim un mínim (pot ser punt d'inflexió) → Això passa quan no hi ha convexitat, no tenim garantia de trobar l'òptim.

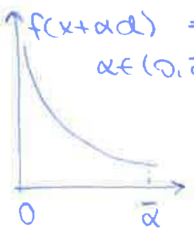
En els mètodes d'optimització que veurem ens anirem movent d'un punt a un altre tq $x^{k+1} = x^k + \alpha^k d^k$. Per a convergir en l'òptim s'ha de complir que d sigui direcció de descens; ha de reduir "localment" el valor de f: $f(x^{k+1}) < f(x^k)$

Definició. Direcció de descens.

$d \in \mathbb{R}^n$ és direcció de descens de $f: \mathbb{R}^n \to \mathbb{R}$ en el punt x^k si $\exists \bar{\alpha} > 0$ tq.

$f(x^k + \alpha d) < f(x^k) \quad \forall \alpha \in (0, \bar{\alpha})$

Gràficament:



$\Rightarrow$ d de descens

Exemple. $f(x_1, x_2) = x_1^2 + x_2^2 \qquad x^k = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

$$x^k + \alpha d = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-\alpha \\ 1-\alpha \end{pmatrix}$$

$$f(x^k + \alpha d) = (1-\alpha)^2 + (1-\alpha)^2 = 2 - 4\alpha + 2\alpha^2 = f(x^k) + \alpha(2\alpha - 4)$$

Per tant, $\forall \alpha \in (0, 2)$, $f(x^k + \alpha d) < f(x^k)$ i d és direcció de descens.

Proposició. Caracterització de les direccions de descens.

Donats $f: \mathbb{R}^n \to \mathbb{R}$ tq $f \in \mathcal{C}^1$, x^k un punt del domini.

Si $\nabla f(x^k)^T \cdot d < 0$, aleshores d és direcció de descens.

Demostració: EXERCICI

⊛ Quantes direccions de descens hi ha en un punt?

Com que $\|\nabla f(x^k)\| \neq 0$ i $\|d\| \neq 0$, tenim que $\nabla f(x^k)^T \cdot d =$

$$= \|\nabla f(x^k)\| \cdot \|d\| \cdot \underbrace{\cos\theta}_{\text{angle entre } \nabla f(x^k),\, d} < 0 \quad \text{si } \cos\theta < 0 \quad \Rightarrow$$

$\Rightarrow$ qualsevol direcció amb $\underbrace{\cos\theta < 0}_{\theta \in (90^\circ, 270^\circ)}$ és de descens $\Rightarrow$ hi ha infinites direccions de descens en un punt.

Mètode del gradient (steepest descent method)

La direcció de descens més ràpid de f al punt x^k és la solució de

$$\min \nabla f(x^k)^T \cdot d$$
$$\text{s.a.: } \|d\|_2 = 1$$

→ s'imposa perquè sinó el problema lineal seria il·limitat.

↳ podem calcular el mínim directament:

$$\nabla f(x^k)^T \cdot d = \|\nabla f(x^k)\| \cdot \underbrace{\|d\|}_{=1} \cdot \underbrace{\cos\theta}_{\geq -1} \geq -\|\nabla f(x^k)\| \Rightarrow$$

$$\Rightarrow d = \frac{-\nabla f(x^k)}{\|\nabla f(x^k)\|}$$ → Aquesta és la direcció de màxim descens.

Utilitzarem $d = -\nabla f(x^k)$, que fa un angle de π rad amb $\nabla f(x^k)$.

Les iteracions del mètode del gradient són:

$$x^{k+1} = x^k + \alpha \nabla f(x^k).$$

⚠ Es tracta d'un mètode generalment lent però que ens garanteix arribar a l'òptim.

Exemple. Funció de Rosenbrock en $\mathbb{R}^2$

$$f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$$

$$\nabla f(x) = \begin{bmatrix} -40(x_2 - x_1^2)x_1 - 2(1 - x_1) \\ 20(x_2 - x_1^2) \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 120x_1^2 - 40x_2 + 2 & -40x_1 \\ -40x_1 & 20 \end{bmatrix}$$

Mínim local estricte: $x^* = \binom{1}{1}$, $\nabla f(x^*) = \binom{0}{0}$, $\nabla^2 f(x^*) = \begin{bmatrix} 82 & -40 \\ -40 & 20 \end{bmatrix}$

↳ és definida positiva

↳ De fet és mínim global encara que f no és convexa.

En el punt $x^k = \binom{1/2}{1/2}$, $f(x^k) = \frac{7}{8}$, $\nabla f(x^k) = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$, $\nabla^2 f(x^k) = \begin{bmatrix} 12 & -20 \\ -20 & 20 \end{bmatrix}$

comprovem condició de descens $\nabla f(x^k)^T \cdot d$ en vàries direccions:

$d^1 = -\nabla f(x^k) = \binom{6}{-5}$; $\nabla f(x^k)^T \cdot d^1 < 0 \Rightarrow d^1$ de descens

$d^2 = \binom{1}{0}$; $\nabla f(x^k)^T \cdot d^2 < 0 \Rightarrow d^2$ de descens

$d^3 = \binom{0}{1}$; $\nabla f(x^k)^T \cdot d^3 > 0 \Rightarrow d^3$ d'ascens

$d^4 = \binom{5}{6}$ (ortogonal a $\nabla f(x^k)$) ; $\nabla f(x^k)^T \cdot d^4 = 0 \Rightarrow$ No sabem si és d'ascens o de descens

↳ prenem $f(x^k + \alpha d^4) = f(x^k) + \alpha \nabla f(x^k)^T \cdot d^4 +$
$+ \frac{1}{2}\alpha^2 (d^4)^T \nabla f(x^k) d^4 + o(\alpha^2)$ (mirem el següent terme del P. Taylor)

Caldria mirar el signe de $(d^4)^T \cdot \nabla^2 f(x^k) \cdot d^4 =$

$$= (5 \quad 6)\begin{pmatrix} 12 & -20 \\ -20 & 20 \end{pmatrix}\begin{pmatrix} 5 \\ 6 \end{pmatrix} < 0 \Rightarrow d^4 \text{ de descens}$$

⊛ La direcció $-\nabla f(x^k)$ és de descens només localment. Si $\alpha >> 0$ aleshores

$f(x^k - \alpha \nabla f(x^k)) > f(x^k)$

una bona α serà suficient a la pràctica

↳ caldria fer una exploració lineal per a trobar la millor α^*

Creem $\begin{aligned} g: \mathbb{R} &\to \mathbb{R} \\ \alpha &\mapsto g(\alpha) \end{aligned} \Big\} \; f(\underbrace{x^k}_{\text{fixat } d} + \alpha(\underline{-\nabla f(x^k)})) = g(\alpha)$

⊛ Algorisme del mètode del Gradient:

punt inicial x^0, $k=0$

while x^k no és solució do

$d^k = -\nabla f(x^k)$

calcular α^k que satisfà condicions A-W

$x^{k+1} := x^k + \alpha^k d^k$

$k := k+1$

end-while

Return: $x^* = x^k$

Mètode de Newton

La direcció de Newton s'obté a partir d'aproximació quadràtica de f al punt x^k: T. Taylor

$$f(x^k + d) \overset{\downarrow}{=} f(x^k) + \nabla f(x^k)^T \cdot d + \frac{1}{2} d^T \cdot \nabla^2 f(x^k) d + o(\|d\|^2)$$

no cal imposar que $\|d\| = 1$ perquè és una funció quadràtica.

La aproximació de f al punt $x^k \equiv m_k(d)$ → model quadràtic de f a x^k

Per tant ens queda: $f(x^k + d) \approx m_k(d) = f(x^k) + \nabla f(x^k)^T \cdot d$

$+ \frac{1}{2} d \; \nabla^2 f(x^k) \; d$

Escrivim aproximació quadràtica en funció de x: $d = x - x^k$

$$m_k(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k)$$

Suposant $\nabla^2 f(x^k) \succcurlyeq 0$ ($m_k(d)$ convexa) calculem

$$\min_{d \in \mathbb{R}^n} m_k(d) = f(x^k) + \nabla f(x^k)^T \cdot d + \frac{1}{2} d^T \nabla^2 f(x^k) d$$

↳ (derivem) $\nabla m_k(d) = \nabla f(x^k) + \nabla^2 f(x^k) \cdot d = 0 \Rightarrow$

perquè estem fent el mínim (pt. estacionari)

$\Rightarrow d = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$ → Aquesta és la direcció de Newton.

⚠ Es tracta d'un mètode ràpid, però a vegades falla.

Exemple. Funció de Rosenbrock en $\mathbb{R}^2$

[mirar l'exemple anterior]

Calculem model quadràtic $m_k(x)$ en $x^k = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$:

$$x^k = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad f(x^k) = 1 \qquad \nabla f(x^k) = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \qquad \nabla^2 f(x^k) = \begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix}$$

$$m_k(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k) =$$

$$= 1 + (-2 \;\; 0)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2}(x_1 \;\; x_2)\begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$

$$= x_1^2 + 10x_2^2 - 2x_1 + 1$$

$\nabla m_k(x) = \begin{bmatrix} 2x_1 - 2 \\ 20x_2 \end{bmatrix}$ el mínim de $m_k(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

La direcció de Newton al punt $x^k = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ és $d = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$

$$\nabla^2 f(x^k)\, d = -\nabla f(x^k) \Rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = -\begin{pmatrix} -2 \\ 0 \end{pmatrix} \Rightarrow d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

El nou punt és $x^{k+1} = x^k + d = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \to$ és el mínim de $m_k(x)$

Calculem model quadràtic $m_{k+1}(x)$ en $x^{k+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$:

$$x^{k+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad f(x^{k+1}) = 10 \qquad \nabla f(x^{k+1}) = \begin{pmatrix} 40 \\ -20 \end{pmatrix} \qquad \nabla^2 f(x^{k+1}) = \begin{pmatrix} 122 & -40 \\ -40 & 20 \end{pmatrix}$$

$$m_{k+1}(x) = f(x^{k+1}) + \nabla f(x^{k+1})^T (x - x^{k+1}) + \frac{1}{2}(x - x^{k+1})^T \nabla^2 f(x^{k+1})(x - x^{k+1}) =$$

$$= 10 + (40 \;\; -20)\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} + \frac{1}{2}(x_1 - 1 \;\; x_2)\begin{pmatrix} 122 & -40 \\ -40 & 20 \end{pmatrix}\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} =$$

$$= 61x_1^2 + 10x_2^2 - 40x_1x_2 - 82x_1 + 20x_2 + 31$$

$\nabla m_{k+1}(x) = \begin{bmatrix} 122x_1 - 40x_2 - 82 \\ 20x_2 - 40x_1 + 20 \end{bmatrix}$ el mínim de $m_{k+1}(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

La direcció de Newton al punt $x^{k+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ és $d = -(\nabla^2 f(x^{k+1}))^{-1} \nabla f(x^{k+1})$

$$\ldots \Rightarrow d = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

El nou punt és $x^{k+2} = x^{k+1} + d = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \to$ és el mínim de $m_{k+1}(x)$

$x^{k+2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ és la solució de $\min f(x)$.

Algorisme del mètode de Newton:

punt inicial x^0, $k=0$

while x^k no és solució do

$d^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$

calcular α^k que satisfà condicions A-W

$x^{k+1} = x^k + \alpha^k d^k$

$k := k+1$

end_while

Return $x^* = x^k$

Propietats de la direcció de Newton:

direcció de newton

1) Si $\nabla^2 f(x) \succ 0 \ \forall x^k$ $\left(\Rightarrow (\nabla^2 f(x))^{-1} \succ 0 \ \forall x^k\right)$ aleshores d és de descens

Demostració

- $\nabla^2 f(x) \succ 0 \Rightarrow (\nabla^2 f(x))^{-1} \succ 0$: EXERCICI
- $\nabla^2 f(x^k) \succ 0 \Rightarrow d$ de descens:

$$\nabla^2 f(x^k) \succ 0 \Rightarrow \underset{\substack{\uparrow \\ d \text{ direcció} \\ \text{de Newton}}}{\nabla f(x^k) \cdot d} = -\underbrace{\nabla f(x^k)^T (\nabla f(x^k))^{-1} \nabla f(x^k)}_{>0} < 0$$

Per tant $\nabla f(x^k) \cdot d < 0 \Rightarrow d$ de descens OK

Fixem-nos que qualsevol B_k tq $B_k \succ 0$ garanteix que $d = -B_k^{-1} \nabla f(x^k)$ de descens

↳ Exemples: si $B_k = Id$ → mètode del gradient

si $B_k = \nabla^2 f(x^k)$ → mètode de Newton

si B_k és aproximació hessiana $(B_k \succ 0)$ → mètode quasi-Newton.

2) A diferència del mètode del gradient, $\alpha = 1$ és el pas "natural" a la direcció de Newton. Si es pot, s'utilitza $\alpha = 1$.

3) Inconvenients:

- si $\nabla^2 f(x^k) \not\succ 0 \Rightarrow$ no podem garantir descens (pot no convergir)
- si $\nabla^2 f(x^k)$ és singular $\Rightarrow$ No té inversa $\Rightarrow$ NO podem calcular d.

↳ cal usar variants que modifiquen la hessiana.

Si estem en un problema convex el mètode de Newton és un bon mètode, altrament no podem assegurar que trobem el mínim.

Exploració lineal exacta i inexacta

Un cop trobada la direcció cap a un punt millor que el que tenim, hem de saber quant ens movem al llarg d'aquella direcció.

És fonamental per a garantir "convergència global".

no va aquí

Definició. Convergència global.

Un mètode d'optimització que genera la seqüència $\{x^k\}_{k\geq 0}$ es diu "globalment convergent" si $\lim_{x\to\infty} x^k = x^*$ $\quad (\nabla f(x^*) = 0) \quad \forall x^0$

↳ és igual per quin punt comencem

L'exploració lineal consisteix a trobar una bona α per a $x^{k+1} = x^k + \alpha d^k$

Definició Exploració lineal exacta.

Parlem d'exploració lineal exacta quan trobem la millor α.

$$\alpha^* = \arg\min_{\alpha\geq 0} g(\alpha) = f(x^k + \alpha d^k)$$

on $\quad g'(\alpha) = \nabla f(x^k + \alpha d^k)^T \cdot d^k \qquad (g' \; \mathbb{R} \to \mathbb{R})$

A la pràctica NO hi ha cap mètode que faci exploració lineal exacta.

Definició. Exploració lineal inexacta.

Com que és computacionalment costós dur a terme l'exploració lineal exacta, a la pràctica n'hi ha prou amb una α subòptima que garanteixi que $\{x^k\}_{k\geq 0}$ convergeix a un punt estacionari.

Quina condició imposem a α per garantir convergència a $\nabla f(x^*) = 0$?

NO és suficient imposar $f(x^k + \alpha d) < f(x^k)$.

↳ Imposem les condicions d'Armijo-Wolfe.

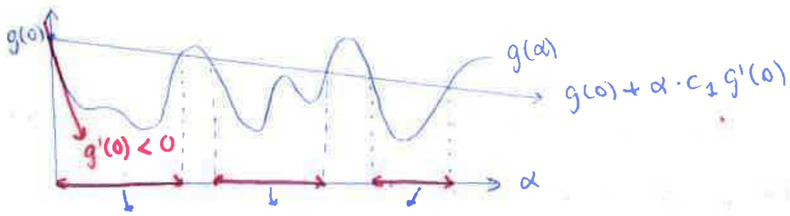
Proposició. Condicions d'Armijo-Wolfe.

1) Condició de descens suficient (AW-1):

$$g(\alpha) \leq g(0) + \alpha c_1 g'(0) \qquad c_1 \in (0,1)$$

$$f(x^k + \alpha d^k) \leq f(x^k) + \alpha c_1 \nabla f(x^k)^T d^k \qquad c_1 \in (0,1)$$

Per exemple:



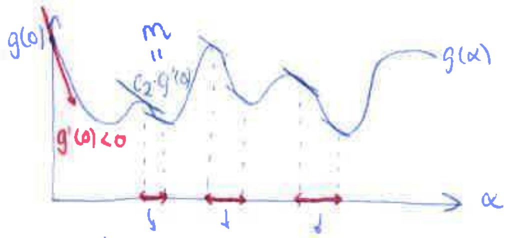
trossos on tenim una bona α.

2) Condició de corbatura (AW-2):

$g'(\alpha) \geq c_2 \cdot g'(0) \qquad 0 < c_1 < c_2 < 1$

$\nabla f(x^k + \alpha d^k)^T \cdot d^k \geq c_2 \nabla f(x^k) d^k \qquad 0 < c_1 < c_2 < 1$

Per exemple:



trossos on tenim una bona α

$c_2 \cdot g'(0) = 0$
Ens serveixen les α on g(α) té un pendent superior a m.

↳ Per a trobar α, necessitem que es compleixin les dues condicions d'Armijo-Wolfe

Convergència global

Si trobem α i d^k que satisfacin unes condicions determinades, aleshores es pot demostrar que la seqüència de punts $x^{k+1} = x^k + \alpha^k d^k$ convergirà globalment a un punt estacionari.

Definició. Convergència global
Un algorisme té convergència global si finalitza en una solució independentment del punt inicial: $\lim_{k \to \infty} x^k = x^* \quad \forall x^0$

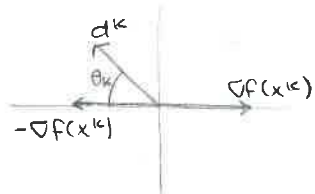
no va aquí

Definició. Convergència local
La convergència local estudia la velocitat a la que ens atansem al punt x^* quan estem a prop d'ell.

⚠ **Notació.**

$\theta_k = \text{angle}(d^k, -\nabla f(x^k))$ (θ_k no pot ser 90° ni 270°)

$$\cos \theta_k = \frac{(-\nabla f(x^k))^T \cdot d^k}{\|\nabla f(x^k)\| \cdot \|d^k\|}$$



Teorema. Teorema de Zoutendijk
Considerem la seqüència $\{x^k\}_{k \geq 0}$ generada per $x^{k+1} = x^k + \alpha^k d^k$, on

1) d^k és direcció de descens
2) α^k satisfà les condicions d'Armijo-Wolfe
   i) $g(\alpha) \leq g(0) + \alpha \cdot c_1 g'(0) \qquad 0 < c_1 < 1$
   ii) $g'(\alpha) \geq c_2 \cdot g'(0) \qquad 0 < c_1 < c_2 < 1$
3) $f: \mathbb{R}^n \to \mathbb{R}$ està fitada inferiorment a $\mathbb{R}^n$
4) $f \in \mathcal{C}^1$ en un entorn obert $\mathcal{N} \supseteq \mathcal{L} = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$ x^0 punt inicial
5) ∇f és Lipschitz continu en $\mathcal{N}$, és a dir, $\exists L \in \mathbb{R}$ tq
   $\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \forall x, y \in \mathcal{N}$

Aleshores, $\sum^{\infty} \cos^2\theta_k \, \|\nabla f(x^k)\|^2 < \infty$.

garanteix la convergència global.

$$\sum^{\infty} \cos^2\theta_k \cdot \|\nabla f(x^k)\|^2 < \infty \Rightarrow \lim_{k\to+\infty} \cos^2\theta_k \cdot \|\nabla f(x^k)\|^2 = 0.$$

Si podem garantir que $\exists\, \delta > 0$ tq. $\forall k$ $\cos\theta_k \geq \delta > 0$, aleshores ($\theta_k < \pi/2$)

$$\lim_{k\to+\infty} \underbrace{\cos^2\theta_k}_{\substack{\vee\!\mid \\ 0}} \cdot \|\nabla f(x^k)\|^2 = 0 \Rightarrow \lim_{k\to+\infty} \|\nabla f(x^k)\|^2 = 0 \Rightarrow$$

$\Rightarrow x^k$ convergeix en un punt estacionari OK.

a) Convergència global del mètode del gradient

$d^k = -\nabla f(x^k)$; per tant $\theta_k = 0$ i $\cos\theta_k = 1$

Per tant, si les longituds de pas α^k verifiquen les condicions AW, el mètode del gradient té convergència global a un punt estacionari.

b) Convergència global del mètode de Newton

$d^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$

Si $\nabla^2 f(x^k)$ és definida positiva d^k és de descens. Suposem que és així.

$\theta_k = ?$

Si $\nabla^2 f(x^k)$ verifica que $\exists\, M < \infty$ tq $\underbrace{\|\nabla^2 f(x^k)\| \cdot \|(\nabla^2 f(x^k))^{-1}\|}_{K(\nabla^2 f(x^k)) \text{ condicionament de la matriu.}} < M \Rightarrow$

$\Rightarrow \cos\theta_k \geq \dfrac{1}{M} > 0$ fita uniforme (no depèn de k)

⚠ Recordatori:

$$\left.\begin{array}{l}\|A\| = \max \|Ax\| \\ \text{s.a.: } \|x\| = 1\end{array}\right\} = \max \frac{\|Ax\|}{\|x\|}$$

Demostració

Si α^k verifica la condició AW-2, $g'(\alpha^k) \geq c_2 \cdot g'(0)$ $\quad 0 < c_1 < c_2 < 1$

Recordem que $g(\alpha) = f(x^k + \alpha d^k)$, $g'(\alpha) = \nabla f(x^k + \alpha d^k)^T \cdot d^k$

$$\nabla f(\underbrace{x^k + \alpha^k d^k}_{x^{k+1}})^T \cdot d^k \geq c_2 \nabla f(x^k)^T \cdot d^k$$

$$\underbrace{(\nabla f(x^{k+1}) - \nabla f(x^k))^T \cdot d^k}_{\downarrow} \geq \overbrace{(c_2 - 1)}^{<0} \overbrace{\nabla f(x^k)^T \cdot d^k}^{<0} > 0 \qquad (1)$$

prenem això, volem trobar una fita superior.

(2) $(\nabla f(x^{k+1}) - \nabla f(x^k))^T \cdot d^k \underset{\substack{\uparrow \\ \text{desigualtat C-S.}}}{\leq} \|\nabla f(x^{k+1}) - \nabla f(x^k)\| \, \|d^k\|$

$$(\nabla f(x^{k+1}) - \nabla f(x^k))^T \cdot d^k \leq L \, \underbrace{\|x^{k+1} - x^k\|}_{\substack{x^{k+1} = x^k + \alpha^k d^k \\ x^{k+1} - x^k = \alpha^k d^k}} \cdot \|d^k\|$$

$$(\nabla f(x^{k+1}) - \nabla f(x^k))^T \cdot d^k \leq L \cdot \alpha^k \|d^k\| \cdot \|d^k\| = L \cdot \alpha^k \|d^k\|^2$$

Usant (1) i (2) tenim:

$$L \alpha^k \|d^k\|^2 \geq (c_2 - 1) \nabla f(x^k)^T \cdot d^k$$

$$\alpha^k \geq \underbrace{\frac{(c_2-1)}{L}}_{<0} \cdot \frac{\overbrace{\nabla f(x^k)^T \cdot d^k}^{<0}}{\|d^k\|^2} > 0 \qquad (3)$$

Usem AW-1: $g(\alpha^k) \leq g(0) + \alpha \cdot c_1 \, g'(0)$

$$f(\underbrace{x^k + \alpha^k d^k}_{x^{k+1}}) \leq f(x^k) + \alpha^k \cdot c_1 \cdot \nabla f(x^k)^T \cdot d^k$$

$$f(x^{k+1}) - f(x^k) \leq \alpha^k \cdot \overset{>0}{c_1} \, \overbrace{\nabla f(x^k)^T \cdot d^k}^{<0} \overset{(3)}{\leq} \underbrace{\frac{(c_2-1)}{L} \cdot c_1}_{= -C} \cdot \frac{(\nabla f(x^k)^T d^k)^2}{\|d^k\|^2} =$$

$$\left(C = \frac{(1-c_2)\, c_1}{L} > 0 \right)$$

$$= -C \cdot \frac{(\nabla f(x^k)^T d^k)^2}{\|d^k\|^2}$$

$$\cos \theta_k = \frac{-\nabla f(x^k)^T \cdot d^k}{\|\nabla f(x^k)\| \cdot \|d^k\|} \; ; \quad \|\nabla f(x^k)\| \cdot \cos \theta_k = \frac{-\nabla f(x^k)^T \cdot d^k}{\|d^k\|} \Rightarrow$$

$$\Rightarrow \|\nabla f(x^k)\|^2 \cdot \cos^2 \theta_k = \frac{(\nabla f(x^k)^T \cdot d^k)^2}{\|d^k\|^2}$$

$$f(x^{k+1}) - f(x^k) \leq -C \cdot \cos^2 \theta_k \cdot \|\nabla f(x^k)\|^2 \qquad (4)$$

Usem (4):

$$f(x^{k+1}) - f(x^k) \leq -C \cdot \cos^2 \theta_k \cdot \|\nabla f(x^k)\|^2$$
$$f(x^k) - f(x^{k-1}) \leq -C \cdot \cos^2 \theta_{k-1} \cdot \|\nabla f(x^{k-1})\|^2$$
$$\vdots$$
$$f(x^1) - f(x^0) \leq -C \cdot \cos^2 \theta_0 \cdot \|\nabla f(x^0)\|^2$$

$$f(x^{k+1}) - f(x^0) \leq -C \cdot \sum_{j=0}^{k} \cos^2 \theta_j \, \|\nabla f(x^0_j)\|^2$$

$$C \cdot \sum_{j=0}^{k} \cos^2 \theta_j \, \|\nabla f(x^0_j)\|^2 \leq f(x^0) - f(x^{k+1}) < T \leftarrow \exists T < \infty$$

(f fitada superiorment)

$$\lim_{k \to \infty} C \cdot \sum_{j=0}^{k} \cos^2 \theta_j \, \|\nabla f(x_j^0)\|^2 \leq \lim_{k \to +\infty} f(x^0) - f(x^{k+1}) < T \Rightarrow$$

$$\Rightarrow \sum_{j=0}^{\infty} \cos \theta_j \, \|\nabla f(x^0_j)\|^2 \leq \frac{T}{C} < \infty$$

<u>OK</u> hem aconseguit una fita superior.

Convergència local i velocitat de convergència

Definició. Convergència local.

La convergència local és l'estudi de la velocitat de convergència quan estem "a prop" de l'òptim.

Com mesurem les velocitats de convergència?

- Convergència lineal: $\{x^k\} \to \{x^*\}$

$$\exists r \in (0,1) \text{ tq. } \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \leq r \quad \forall k \text{ "suficientment" gran.}$$

(mètode del gradient) → lenta.

- Convergència superlineal: cas límit de la velocitat de convergència lineal.

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0$$

velocitat intermitja

- Convergència quadràtica:

$$\exists M \in \mathbb{R} \text{ tq } \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \leq M \quad \forall k \text{ "suficientment" gran}$$

(mètode de Newton) → ràpida.

1) Convergència local del mètode del gradient

a) Funcions quadràtiques:

tenim $\min_{x \in \mathbb{R}^n} f(x)$, on $f(x) = \frac{1}{2} x^T Q x - b^T x$ $\left(\begin{array}{l} Q \text{ simètrica} \\ Q \text{ definida positiva} \end{array} \right)$ → f quadràtica

$\nabla f(x) = Qx - b$ → si imposem $\nabla f(x) = 0$ com que $Qx = b$ $\}$ $\nabla f(x) = Qx - b$

$\nabla^2 f(x) = Q \succ 0$

↓

Aleshores $x^{k+1} = x^k - \alpha^k \nabla f(x^k)$ $\quad (d = -\nabla f(x^k))$

amb el mètode del gradient podem resoldre sistemes d'aquest tipus.

Propietats:

- Podem calcular analíticament α^k per exploració lineal exacta:

$$\alpha^k = \arg\min_{\alpha > 0} f(x - \alpha \nabla f(x^k)) \Rightarrow \alpha^k = \frac{\nabla f(x^k)^T \nabla f(x^k)}{\nabla f(x^k)^T Q \nabla f(x^k)} > 0$$

- Si es fa exploració lineal exacta, dues direccions consecutives són ortogonals.

$(d^k)^T \cdot d^{k+1} = 0$

↑ això és cert per a $f: \mathbb{R}^n \to \mathbb{R}$ en general.

⚠ Recordem:

$\|x\|_Q^2 = x^T \cdot Q \cdot x$ (norma ponderada per Q simètrica i def pos)

Q def pos $\Rightarrow$ VAPS $0 < \lambda_1 \leq \ldots \leq \lambda_n$

Teorema. Velocitat de convergència del mètode del gradient.

Si apliquem el mètode del gradient a

i) una funció quadràtica $f(x) = \frac{1}{2} x^T Q x - b^T x$, $Q \succ 0$ i simètrica

ii) amb exploració lineal exacta $x^k = \arg\min_{\alpha > 0} f(x^k + \alpha d^k)$

aleshores: $\dfrac{f(x^{k+1}) - f(x^*)}{f(x^k) - f(x^*)} = \dfrac{\|x^{k+1} - x^k\|_Q^2}{\|x^k - x^*\|_Q^2} \leq \left(\dfrac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2$ on λ_1 = vap més petit de Q, λ_n = vap més gran de Q

Per a demostrar-ho cal la desigualtat de Kantorovich

Aquest teorema ens diu que:

- si $\lambda_1 \approx \lambda_n \Rightarrow \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \approx 0 \longrightarrow$ serà ràpid
- si $\lambda_n >> \lambda_1 \Rightarrow \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 >> 0 \longrightarrow$ serà lent

Exemple. $f(x) = \frac{1}{2} x^T \cdot Q \cdot x - b^T x$

- $Q = \begin{pmatrix} 20 & 4 \\ 4 & 1 \end{pmatrix}$ $b = \begin{pmatrix} 24 \\ 5 \end{pmatrix}$ $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\left(\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1}\right)^2 = 0'964 \rightarrow$ convergència en 118 iteracions

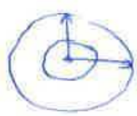
  $\lambda_1 = 0'192$, $\lambda_2 = 20'808$

- $Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ $b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\left(\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1}\right)^2 = 0'555 \rightarrow$ convergència en 32 iteracions

  $\lambda_1 = 0'382$, $\lambda_2 = 2'618$

En les formes quadràtiques, les corbes de nivell són el·lipses. Els vaps de la forma quadràtica ens indiquen la relació dels eixos de la el·lipse:

vaps molt diferents



vaps molt semblants

b) Funcions generals $f: \mathbb{R}^n \rightarrow \mathbb{R}$

Teorema. Suposem que el mètode del gradient

i) amb exploració lineal exacta $x^k = \arg\min_{\alpha > 0} f(x^k + \alpha d^k)$

ii) aplicat a $f: \mathbb{R}^n \rightarrow \mathbb{R}$

iii) convergeix a un punt x^* on $\nabla^2 f(x^*)$ és definida positiva.

Aleshores, per a k prou gran

$$\frac{f(x^{k+1})-f(x^k)}{f(x^k)-f(x^*)} \leq r^2 \quad \text{on} \quad r \in \left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}, 1\right) \quad \text{on} \quad \begin{array}{l}\lambda_1 = \text{vap més petit de } \nabla^2 f(x^*) \\ \lambda_n = \text{vap més gran de } \nabla^2 f(x^*)\end{array}$$

2) Convergència local del mètode de Newton

A mida que ens apropem a x^* pot demostrar-se que la longitud de pas $\alpha^k = 1$ verificarà les condicions d'Armijo-Wolfe. Per tant podem suposar que a prop de l'òptim (o per k suficientment gran) les iteracions del mètode de Newton seran:

$$x^{k+1} = x^k + d^k \qquad d^k = -(\nabla^2 f(x^k))^{-1}\nabla f(x^k) \quad (\text{suposem } d^k \text{ de descens})$$

Teorema. Convergència local del mètode de Newton.

Suposem que:

1) $f \in \mathcal{C}^2$

2) $\nabla^2 f$ és Lipschitz contínua en un entorn $\mathcal{N}$ de x^*, és a dir,
$\exists L > 0$ tq $\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x-y\| \quad \forall x,y \in \mathcal{N}$

3) A x^* es verifiquen les condicions suficients d'optimalitat

 i) $\nabla f(x^*) = 0$

 ii) $\nabla^2 f(x^*) > 0$

4) Considerem la iteració $x^{k+1} = x^k + d^k$, on $d^k = -(\nabla^2 f(x^k))^{-1}\nabla f(x^k)$, $\alpha^k = 1$.

Aleshores:

i) Si x^0 és "prou proper" a x^* aleshores $\{x^k\} \xrightarrow[k\to\infty]{} x^*$

ii) la seqüència $\{x^k\}_{k\geq 0}$ convergeix a x^* amb velocitat quadràtica

iii) la seqüència $\{\|\nabla f(x^k)\|\} \xrightarrow[k\to\infty]{} 0$ amb velocitat quadràtica

<u>Demostració</u>

ii) $\{x^k\}_{k\geq 0}$ convergeix quadràticament.

Hem de provar que $\exists M \in \mathbb{R}^+$ tq $\dfrac{\|x^{k+1}-x^*\|}{\|x^k - x^*\|^2} \leq M$

$$x^{k+1} = x^k + d^k$$

$$x^{k+1} - x^* = x^k + d^k - x^* = x^k - x^* - (\nabla^2 f(x^k))^{-1}\nabla f(x^k) =$$

$$= \underbrace{(\nabla^2 f(x^k))^{-1}}_{(a)} \underbrace{\left[\nabla^2 f(x^k)(x^k - x^*) - (\nabla f(x^k) - \underbrace{\nabla f(x^*)}_{=0})\right]}_{(b)}$$

L'objectiu de la nostra demostració és fitar la norma de (a) i de (b).

b) $\| \nabla^2 f(x^k)(x^k - x^*) - (\nabla f(x^k) - \nabla f(x^*)) \| =$

$$= \| \nabla^2 f(x^k)(x^k - x^*) - \int_0^1 \nabla^2 f(x^k + t(x^* - x^k))\underbrace{(x^k - x^*)}\, dt \| =$$

hi ha un canvi de signe perquè tenim la diferència dels dos punts al revés.

considerem $\Phi : \mathbb{R} \to \mathbb{R}^n$

$\Phi(t) = \nabla f(x^k + t(x^* - x^k))$

$\Phi(0) = \nabla f(x^k)$; $\Phi(1) = \nabla f(x^*) = 0$

$\Phi'(t) = \nabla^2 f(x^k + t(x^* - x^k))(x^* - x^k)$

Apliquem T.F.C. : $\Phi(1) - \Phi(0) = \int_0^1 \Phi'(t)\, dt$

$$\nabla f(x^*) - \nabla f(x^k) = \int_0^1 \nabla^2 f(x^k + t(x^* - x^k))(x^* - x^k)\, dt$$

$$= \| \int_0^1 (\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k)))(x^k - x^*)\, dt \| \leq$$

$$\leq \int_0^1 \| (\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k)))\overbrace{(x^k - x^*)}^{\text{recta}} \|\, dt \leq$$

desigualtat triangular : $\langle u, v \rangle = u^T \cdot v = \sum_{i=1}^n u_i v_i$

$\langle f, g \rangle = \int_{t \in T} f(t) \cdot g(t)\, dt$; $\|f\|^2 = \int_{t \in T} (f(t))^2\, dt$

$\Rightarrow \|u + v\| \leq \|u\| + \|v\|$

$\| \sum_{i=1}^n v_i \| \leq \sum_{i=1}^n \|v_i\| \rightarrow \| \int f(t)\, dt \| \leq \int \|f(t)\|\, dt$ (la norma de la integral és més petita que la integral de la norma)

$$\leq \int_0^1 \| \nabla^2 f(x^*) - \nabla^2 f(x^k + t(x^* - x^k)) \| \cdot \| x^k - x^* \|\, dt \leq$$

propietats norma matriu: $\|A\| = \max \frac{\|Ax\|}{\|x\|}$; $\frac{\|Ax\|}{\|x\|} \leq \|A\| \Rightarrow \|A\| \cdot \|x\| \geq \|Ax\|$

$$\leq \int_0^1 L \| \cancel{x^k} + t(x^* - x^k) - \cancel{x^*} \| \cdot \| x^k - x^* \|\, dt =$$

Lipschitz

$$= \int_0^1 L t \| x^k - x^* \|^2\, dt = \| x^k - x^* \|^2 \cdot L \int_0^1 t\, dt = \boxed{\frac{L}{2} \| x^k - x^* \|^2}$$

a) $\nabla^2 f(x^*) \succ 0 \Rightarrow \exists (\nabla^2 f(x^*))^{-1}$

$0 < \| (\nabla^2 f(x^*))^{-1} \| < \infty$

$\exists r \mid \forall x^k \quad \| x^k - x^* \| < r$

$\| (\nabla^2 f(x^k))^{-1} \| < \boxed{2 \| (\nabla^2 f(x^k))^{-1} \|}$

Així doncs ja tenim les dues normes fitades.

$$\|x^{k+1}-x^k\| = \dots = \|(\nabla^2 f(x^k))^{-1}(\nabla^2 f(x^k)(x^k-x^*) - (\nabla f(x^k)-\nabla f(x^*))\|$$

Per tant, $\|x^{k+1}-x^k\| = \|(\nabla^2 f(x^k))^{-1}(\nabla^2 f(x^k)(x^k-x^*)-(\nabla f(x^k)-\nabla f(x^*))\| \leq$

$$\leq \|(\nabla^2 f(x^k))^{-1}\| \cdot \|\nabla^2 f(x^k)(x^k-x^*) - (\nabla f(x^k) - \nabla f(x^*))\| \leq$$

$$\leq 2\|(\nabla^2 f(x^*))^{-1}\| \cdot \frac{L}{2}\|x^k-x^*\|^2$$

$$\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|^2} \leq L \cdot \underbrace{\|(\nabla^2 f(x^*))^{-1}\|}_{M} \quad \text{OK}$$

5.- Modificacions al mètode de Newton

Modificació de la Hessiana

Si $\nabla^2 f(x^k)$ no és definida positiva $d^k = -(\nabla^2 f(x^k))^{-1} \cdot \nabla f(x^k)$ no té per què ser de descens (ni existir) i no garantiríem les condicions del teorema de Zoutendijk: no podem garantir convergència global.

Al mètode de Newton modificat usem B_k si $\nabla^2 f(x^k)$ no és prou def. positiva.

$B_k = \nabla^2 f(x^k) + E_k$ on $E_k = 0$ si $\nabla^2 f(x^k)$ és prou def. pos, sinó E_k s'escull de forma que B_k sigui prou def pos.

Convergència global i local del mètode de Newton modificat

- convergència global:

  Si B_k té un nombre de condició uniformement afitat, és a dir,

  $\exists\, C$ tq. $\|B_k\| \cdot \|B_k^{-1}\| \leq C \quad \forall k$

  (és a dir, B_k no és singular ni propera a ser-ho)

  aleshores es demostra que $\cos\theta_k \geq \frac{1}{C}$ i pel T. Zoutendijk

  $\lim_{k\to+\infty} \|\nabla f(x^k)\| = 0 \Rightarrow$ el mètode de Newton modificat té convergència global.

- convergència local:
  - Si $\{x^k\} \to x^*$ i $\nabla^2 f(x^k)$ és prou def pos ($E_k = 0$ a partir d'un k' determinat), aleshores el mètode de Newton modificat es comporta com el mètode de Newton i té convergència local quadràtica.
  - Si $\{x^k\} \to x^*$ i $\nabla^2 f(x^*)$ no és prou def pos i és gairebé singular, ($E_k \neq 0 \ \forall k$) aleshores la velocitat de convergència pot ser només lineal.

Tipus de modificacions

1) Modificació de la descomposició espectral:

$\nabla^2 f(x^k) = V \Lambda V^T$, $V = [v_1 | \dots | v_n]$ matriu de veps

$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ matriu de vaps

si $\nabla^2 f(x^k)$ no és prou def pos, algun $\lambda_i < \varepsilon$ i el modifiquem $\lambda_i + \delta_i$:

$B_k = V(\Lambda + \Delta) V^T = V\Lambda V^T + V \Delta V^T = \nabla^2 f(x^k) + E_k$

⊛ Inconvenient: cal fer la descomposició espectral de $\nabla^2 f(x^k)$ → costós

2) Modificació per addició d'una matriu diagonal:

$B_k = \nabla^2 f(x^k) + \tau \cdot I$. Si τ és prou gran, aleshores B_k és prou def pos. OK

⊛ Inconvenient: no sabem el valor de τ; es pot anar modificant iterativament fins que es prova (Cholesky) que B_k és prou def. pos → costós

3) Modificació de la factorització de Cholesky:

Factorització de Cholesky: $\nabla^2 f(x^k) = L \cdot D \cdot L^T$, on $D = \text{diag}(d_1, \dots, d_n)$.

Si $\nabla^2 f(x^k)$ no és prou def. pos, algun $d_i < \varepsilon$

Aleshores es modifiquen (s'incrementen) els $d_i < \varepsilon$ durant la factorització de Cholesky.

Finalment tenim una factorització: $LDL^T = P \cdot \nabla^2 f(x^k) \cdot P^T + E_k$

↳ És un mètode eficient.

Exemple. $f(x_1, x_2) = x_1 \cdot e^{-x_1^2 - x_2^2}$

$\nabla f(x) = e^{-x_1^2 - x_2^2} \begin{bmatrix} 1 - 2x_1^2 \\ -2x_1 x_2 \end{bmatrix}$; $\nabla^2 f(x) = e^{-x_1^2 - x_2^2} \begin{bmatrix} -6x_1 + 4x_1^3 & -2x_2 + 4x_1^2 x_2 \\ -2x_2 + 4x_1^2 x_2 & -2x_1 + 4x_1 x_2^2 \end{bmatrix}$

↳ solució usant condicions d'optimalitat:

$\nabla f(x) = 0 \Rightarrow x^+ = \begin{bmatrix} \sqrt{1/2} \\ 0 \end{bmatrix}$

$x^- = \begin{bmatrix} -\sqrt{1/2} \\ 0 \end{bmatrix} \rightarrow \nabla^2 f(x^-) = e^{-1/2} \begin{bmatrix} 4\sqrt{1/2} & 0 \\ 0 & 2\sqrt{1/2} \end{bmatrix}$

$\nabla^2 f(x^-)$ és def. pos. ; x^- és un mínim. (x^+ és un màxim)

- Considerem mètode de Newton a partir de $x^0 = \binom{0}{0}$

$\nabla f(x^0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $\nabla^2 f(x^0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \nexists\, d^0$ tq $\nabla^2 f(x^0) \cdot d^0 = -\nabla f(x^0)$

La direcció de Newton no està definida. Mètode de Newton aborta.

- Si considerem el mètode de Newton modificat a partir de x^0, trobem l'òptim a la 5a iteració (amb modificació de la factorització de Cholesky)

TEMA 6 : Optimització no lineal amb restriccions

1.- Motivació de les condicions d'optimalitat

Definició. Problema d'optimització no lineal amb restriccions

$$\begin{aligned} &\min f(x) \\ &\text{s.a.: } h(x) = 0 \qquad [\, h_i(x) = 0 \quad i = 1, \dots, m \,] \\ &\qquad\;\; g(x) \leq 0 \qquad [\, g_j(x) \leq 0 \quad j = 1, \dots, p \,] \end{aligned}$$

on $f: \mathbb{R}^n \to \mathbb{R}$ és la funció objectiu (suposem suau, $f \in \mathcal{C}^2$)

$h: \mathbb{R}^n \to \mathbb{R}^m$ ($h_i \in \mathcal{C}^2$, $i = 1, \dots, m$)

$g: \mathbb{R}^n \to \mathbb{R}^p$ ($g_j \in \mathcal{C}^2$, $j = 1, \dots, p$)

Aleshores el problema $\left.\begin{aligned} &\min f(x) \\ &\text{s.a.: } h(x) = 0 \\ &\qquad g(x) \leq 0 \end{aligned}\right\}$ és un problema d'ONL amb restriccions

Definició. Restriccions actives i inactives

Diem que una restricció $g_j(x) \leq 0$ és activa al punt factible x^1 si $g_j(x^1) = 0$, i diem que és inactiva en aquell punt si $g_j(x^1) < 0$.

Les restriccions $h(x) = 0$ són sempre actives en un punt factible.

Denotem el conjunt de restriccions de desigualtat actives en un punt factible x com $A(x) = \{ j \in \{1, \dots, p\} \mid g_j(x) = 0 \}$

⚠ **El problema que tenim amb la ONL amb restriccions és que en l'òptim no sabem quines restriccions són actives i quines inactives. Sinó, se'ns facilitaria molt la feina.**

Exemples.

1) Una restricció d'igualtat:

$$\begin{aligned} &\min\ x_1 + x_2 \rightarrow f(x) \\ &\text{s.a.: } x_1^2 + x_2^2 - 2 = 0 \rightarrow h(x) \end{aligned}$$

$\nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $\nabla h(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$

El mínim és $x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \Rightarrow \nabla h(x^*) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$

↳ veurem que en qualsevol punt que no sigui x^* podem moure'ns per la regió factible i disminuir f.

$x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ verifica que $\nabla f(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ és paral·lel a $\nabla h(x^*) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$

$\nabla f(x^*) = -\lambda^* \nabla h(x^*) \iff \nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$ — λ és el multiplicador de Lagrange.

En el nostre cas, $\lambda^* = 1/2$

La condició anterior es deriva usant sèries de Taylor de 1r ordre.

Suposem que estem a x factible ($h(x) = 0$) i ens movem al llarg de la direcció d.

d ha de ser

- factible: $h(x+d) = 0$
  $\underbrace{h(x+d)}_{} = h(x) + \nabla h(x)^T \cdot d + o(\|d\|)$
  la condició per a què això sigui factible és $\boxed{\nabla h(x)^T \cdot d = 0}$
- de descens: volem que $\boxed{\nabla f(x)^T \cdot d < 0}$ (condició de descens)

És a dir, si en un punt tenim que $\left.\begin{matrix} \nabla h(x)^T \cdot d = 0 \\ \nabla f(x)^T \cdot d < 0 \end{matrix}\right\}$ llavors x no és l'òptim i podem millorar el valor de la funció objectiu.

Prenem una d factible. (com és que això no pot

Casos $\begin{cases} \text{si } \nabla f(x)^T \cdot d < 0 & \Rightarrow x \text{ no és l'òptim} \\ \text{si } \nabla f(x)^T \cdot d > 0 & \Rightarrow \text{prenem } d := -d \text{, aleshores serà } < 0 \text{ i } x \text{ NO és l'òptim} \\ \text{si } \nabla f(x)^T \cdot d = 0 & \Rightarrow \forall d \text{ factible } \nabla f(x) = 0 \Rightarrow x \text{ és l'òptim} \end{cases}$

(Qualsevol vector $\perp$ a $\nabla h(x)$ ha de ser $\perp$ a $\nabla f(x)$ $\Rightarrow$ ∇f, ∇h colineals) ?

Definició. Funció lagrangiana.

La funció lagrangiana és $\mathcal{L}(x, \lambda) = f(x) + \lambda h(x)$.

⊛ La condició d'optimalitat $\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$ és equivalent a dir que en la solució x^* hi ha un multiplicador de Lagrange tal que $\boxed{\nabla_x \mathcal{L}(x^*, \lambda^*) = 0}$ → El gradient respecte x de la lagrangiana ha de ser zero.

↳ Aquesta condició és necessària, no suficient. En l'exemple anterior, si prenem el punt $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ es satisfà la condició d'optimalitat per a $\lambda = -1/2$, però x no és un mínim sinó un màxim.

⚠ Mirant el signe de λ NO podem identificar si el punt en què estem és max o bé min.

2) Una restricció de desigualtat:

$$\min\ x_1 + x_2 \quad \rightarrow f(x)$$
$$\text{s.a.: } x_1^2 + x_2^2 - 2 \leq 0 \quad \rightarrow g(x)$$

$\nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $\nabla g(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$

Com a l'exemple 1), el mínim és al punt $x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ i es verifica que $\nabla f(x^*) + \mu^* \nabla g(x^*) = 0$ per a x^* i $\mu = 1/2$.

↳ A diferència de les restriccions d'igualtat, aquí el signe del multiplicador de Lagrange μ^* sí que és rellevant.

Un punt x^* és òptim si no hi ha cap direcció factible i de descens.

- d factible: $g(x+d) \leq 0 \Rightarrow g(x) + \nabla g(x)^T \cdot d \leq 0$ (condició de factibilitat) 1r ordre
- d de descens: $f(x+d) < f(x) \Rightarrow \cancel{f(x)} + \nabla f(x)^T \cdot d < \cancel{f(x)} \Rightarrow$

$\Rightarrow \boxed{\nabla f(x)^T \cdot d < 0}$ OK. $\Rightarrow \nabla f(x) \neq 0$ (condició de descens) 1r ordre

com que podem canviar d per $-d$?

Considerem la restricció $g(x) \leq 0$. Sigui d una direcció factible, aleshores $g(x+d) \leq 0 \Rightarrow g(x) + \nabla g(x)^T \cdot d \leq 0$

Pot ser que: 1.- $g(x) = 0$ (restricció activa)

si $\nabla g(x)^T \cdot d \leq 0$ → defineix un semiplà tancat
si $\nabla f(x)^T \cdot d < 0$ → defineix un semiplà obert } →

→ si en el punt x hi ha una direcció d que pertany a la intersecció dels dos semiplans, aleshores x NO és l'òptim

2.- $g(x) < 0$ (restricció inactiva)

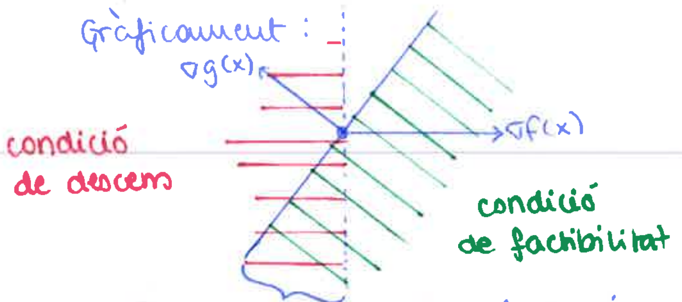
$\underbrace{g(x)}_{\hat{}\ 0} + \nabla g(x)^T \cdot d \leq 0 \Rightarrow$ qualsevol d és factible. (si té una longitud suficientment petita).

Per a que d sigui de descens s'ha de complir $\nabla f(x)^T \cdot d < 0$

(Prenem $d = -\alpha \nabla f(x) \Rightarrow d$ de descens.)

Sempre que $\nabla f(x) \neq 0$ $\exists d$ tq d de descens.

Gràficament:



intersecció dels dos semiplans → NO estem en el punt òptim

⊗ Quan $\nabla g(x)$ sigui $-\nabla f(x)$ NO trobarem cap d que compleixi factibilitat i descens alhora. $\nabla f(x) + \mu \cdot \nabla g(x) = 0$ → Aleshores si $\mu \geq 0$ pot ser que estiguem en l'òptim

↳ $\nabla f(x^*) = -\mu^* \cdot \nabla g(x^*)$
$\mu^* \geq 0$ ← IMPORTANT EL SIGNE

Així doncs, els dos casos anteriors es resumeixen en les condicions següents:

i) $\nabla_x \mathcal{L}(x^*, \mu^*) = \nabla f(x^*) + \mu^* \nabla g(x^*) = 0$

ii) $\mu^* \geq 0$

iii) $\underbrace{\mu^*}_{\geq 0} \underbrace{g(x^*)}_{\leq 0} = 0$ (condició de complementarietat)

↳ si $g(x^*) < 0$ (inactiva) $\Rightarrow \mu^* = 0 \Rightarrow \nabla f(x^*) = 0$. → mateixa condició que en problemes sense restriccions

↳ si $g(x^*) = 0$ (activa) $\Rightarrow \mu^* \geq 0 \Rightarrow \underbrace{\nabla f(x^*)}_{=0} + \underbrace{\mu^* \nabla g(x^*)}_{=0} = 0$

3) Dues restriccions de desigualtat:

$\min \; x_1 + x_2$ → $f(x)$

s.a.: $x_1^2 + x_2^2 - 2 \leq 0$ → $g_1(x)$

$-x_2 \leq 0$ → $g_2(x)$

recomenable posar sempre les restriccions de desigualtat ≤ 0 (si són ≥ 0 el signe de μ^* ha de ser ≤ 0)

Tenim que $\nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$, $\nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

El mínim és $x^* = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}$ i les dues restriccions són actives en x^*.

↳ un punt NO és solució si hi ha una direcció factible i de descens.

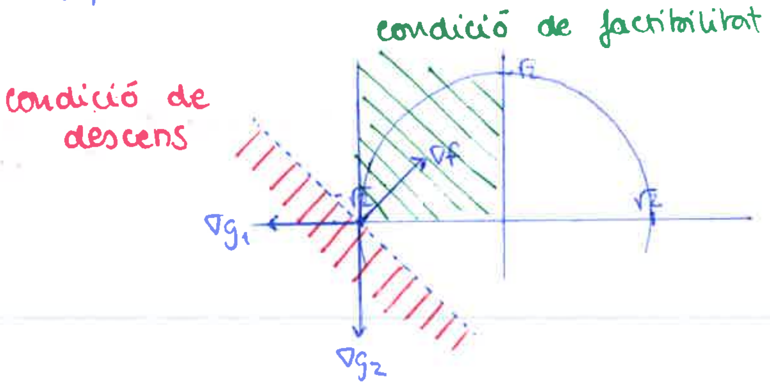
- mantenir factibilitat: $g(x) \leq 0$, $g(x+d) \leq 0 \Rightarrow$

$\Rightarrow \underbrace{g(x)}_{\leq 0} + \nabla g(x)^T \cdot d \leq 0 \overset{?}{\Rightarrow} \nabla g(x)^T \cdot d \leq 0$

NO ENTENC AQUESTA IMPLICACIÓ (*) [illegible] als apunts, no al power)

- Direcció de descens: $\nabla f(x)^T \cdot d < 0$

Gràficament, veiem que no hi ha intersecció dels semiplans en el punt $x^* = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}$:



No hi ha cap direcció que verifiqui factibilitat i descens.

Geomètricament, la condició anterior equival a dir que $-\nabla f(x^*) \in C$, on C és el con format pels gradients de les restriccions actives a x^*.

$$C = \{ v \in \mathbb{R}^n \mid v = \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*), \; \mu_j \geq 0 \}$$

($C \subseteq \mathbb{R}^n$ és un con si $\forall x \in C$ és té que $\alpha x \in C$, on $\alpha \geq 0$, $\alpha \in \mathbb{R}$)

Les condicions anteriors són:

$$-\nabla f(x^*) = \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*) \qquad \mu_1^*, \mu_2^* \geq 0$$

Usant la lagrangiana tenim les condicions:

i) $\nabla_x \mathcal{L}(x^*, \mu^*) = \nabla f(x^*) + \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*) = 0$

ii) $\mu^* = \begin{pmatrix} \mu_1^* \\ \mu_2^* \end{pmatrix} \geq 0$

iii) $\mu_1^* g_1(x^*) = \mu_2^* g_2(x^*) = 0$ (condició de complementarietat)

Condicions KKT (són necessàries de 1r ordre) $\Downarrow$ No ens garanteixen que estiguem en un mínim.

- $x = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}$; $\nabla g_1(x) = \begin{pmatrix} -2\sqrt{2} \\ 0 \end{pmatrix}$; $\nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

(òptim)

$\hookrightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_1^* \begin{pmatrix} -2\sqrt{2} \\ 0 \end{pmatrix} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \mu^* = \begin{pmatrix} \frac{1}{2\sqrt{2}} \\ 1 \end{pmatrix} \geq 0$ OK

- $x = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$ g_1 i g_2 són actives

⚠ $-\nabla f(x) \notin$ con generat pels gradients de les restriccions actives: hi ha direccions factibles i de descens (per exemple $d = (0, -1)^T$)

$\hookrightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_1^* \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \mu^* = \begin{pmatrix} -\frac{1}{2\sqrt{2}} \\ 1 \end{pmatrix}$ $\not\geq 0$

no compleix condicions necessàries d'optimalitat.

- $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ g_1 inactiva i g_2 activa.

⚠ $-\nabla f(x) \notin$ con generat per $\nabla g_1(x)$: hi ha direccions factibles i de descens (per exemple $d = (-1/2, 1/4)^T$

$\hookrightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \cancel{\mu_1^* \begin{pmatrix} 0 \\ 0 \end{pmatrix}} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \nexists \mu_2^*$

2.- Condicions de segon ordre.

Considerarem l'exemple 1) anterior per a derivar condicions de 2n ordre:

$$\left.\begin{array}{l} \min\ x_1 + x_2 \\ \text{s.a.: } x_1^2 + x_2^2 - 2 = 0 \end{array}\right\} \qquad x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

La derivació és: $\mathcal{L}(x^*, \lambda^*) = f(x^*) + \lambda^* h(x^*)$ on $h(x^*) = 0 \Rightarrow$

$\Rightarrow \mathcal{L}(x^*, \lambda^*) = f(x^*)$

⊛ Condicions de 1r ordre: $\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$.

Considerem direcció factible a x^* d:

$h(x^* + d) = 0$ i $\nabla h(x^*)^T \cdot d = 0$

Aleshores, $\mathcal{L}(x^* + d, \lambda^*) = f(x^* + d) + \lambda^* \overbrace{h(x^* + d)}^{0} = f(x^* + d)$

Expansió de la sèrie de Taylor de $\mathcal{L}(x^*+d, \lambda^*)$ en x^*:

$$f(x^*+d) = \mathcal{L}(x^*+d, \lambda^*)$$
$$= \mathcal{L}(x^*, \lambda^*) + \nabla_x \mathcal{L}(x^*, \lambda^*)^T \cdot d + \frac{1}{2} d^T \cdot \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) d + o(\|d\|_2^2)$$
$$= f(x^*) + \frac{1}{2} d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) d + o(\|d\|_2^2)$$

Utilitzant l'expressió anterior i fent tendir $\|d\| \to 0$, aleshores:

- Si x^* és mínim, és necessari que:

$$d^T \nabla^2_{xx} \mathcal{L}(x^*, d^*) d \geq 0 \quad \forall d \mid \nabla h(x^*)^T \cdot d = 0$$

- És suficient per a que x^* sigui mínim que:

$$d^T \nabla^2_{xx} \mathcal{L}(x^*, d^*) \cdot d > 0 \quad \forall d \mid \nabla h(x^*)^T \cdot d = 0$$

Exemple. (està fet de nou més endavant).

multiplicadors de Lagrange.

$$\min \; x_1^2 + x_2^2 \to f(x)$$

s.a.: ($a > 0$)

$$x_1 + x_2 - a = 0 \to h(x) \Rightarrow \lambda$$
$$-x_1 \leq 0 \to g_1(x) \Rightarrow \mu_1$$
$$-x_2 \leq 0 \to g_2(x) \Rightarrow \mu_2$$

Escrivim la lagrangiana:

$$\mathcal{L}(x_1, x_2, \lambda, \mu_1, \mu_2) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - a) + \mu_1(-x_1) + \mu_2(-x_2)$$

Condicions KKT:

i) $x_1 + x_2 = a \qquad x_1, x_2 \geq 0$ <u>OK</u> → factibilitat.

ii) $\nabla_x \mathcal{L}(x, \lambda, \mu) = 0 \Rightarrow \begin{cases} 2x_1 + \lambda - \mu_1 = 0 \\ 2x_2 + \lambda - \mu_2 = 0 \end{cases}$

iii) $\mu_1, \mu_2 \geq 0$

iv) $\mu_1 \cdot x_1 \geq 0, \quad \mu_2 \cdot x_2 \geq 0$

<u>Casos</u>

a) 2 restriccions actives $\Rightarrow x_1 = 0, \; x_2 = 0$

No es compleix la factibilitat $x_1 + x_2 = a \; (a > 0) \Rightarrow$ <u>NO</u>.

b) 1 restricció activa $\Rightarrow$ $x_1 = 0$, $x_2 > 0$

Aleshores, com que $\mu_2 \cdot x_2 = 0 \Rightarrow \mu_2 = 0$ $(\mu_1 \geq 0)$

$\cancel{x_1} + x_2 - a = 0 \Rightarrow x_2 = a$

Seguint amb les condicions,
$$\left.\begin{array}{l} 2\cancel{x_1} + \lambda - \mu_1 = 0 \\ 2x_2 + \lambda - \cancel{\mu_2} = 0 \Rightarrow 2a + \lambda = 0 \end{array}\right\}$$

$$\left.\begin{array}{l} \lambda = \mu_1 \\ \lambda = -2a \end{array}\right\} \Rightarrow \mu_1 = -2a < 0 \Rightarrow \underline{\text{NO}}$$

c) 1 restricció activa $\Rightarrow$ $x_2 = 0$, $x_1 > 0$

↳ NO (per simetria amb el cas b))

d) 2 restriccions inactives $\Rightarrow$ $x_1 > 0$, $x_2 > 0$

Aleshores, com que $\mu_1 x_1 = 0 \Rightarrow \mu_1 = 0$
$\mu_2 x_2 = 0 \Rightarrow \mu_2 = 0$

seguint amb les condicions,
$$\left.\begin{array}{l} 2x_1 + \lambda + \cancel{\mu_1} = 0 \\ 2x_2 + \lambda + \cancel{\mu_2} = 0 \end{array}\right\}$$

$$\left.\begin{array}{l} 2x_1 + \lambda = 0 \\ 2x_2 + \lambda = 0 \\ x_1 + x_2 - a = 0 \end{array}\right\} \begin{array}{l} -2x_1 = -2x_2 \Rightarrow x_1 = x_2 \\ 2x_1 - a = 0 \Rightarrow \boxed{x_1 = a/2} \\ \cancel{2}\frac{a}{\cancel{2}} + \lambda = 0 \Rightarrow \boxed{\lambda = -a} \end{array}$$

Hem de mirar les condicions de 2n ordre:

$\nabla_{xx} \mathcal{L}(x, \lambda, \mu) \succ 0 \Rightarrow$ condició suficient. $\Rightarrow$ OK

Exemple.

$$\left.\begin{array}{l} \min\ x_1 + x_2 \to f(x) \\ \text{s.a.: } x_1^2 + x_2^2 - 2 = 0 \to h(x) \end{array}\right\} \qquad x^* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\nabla f(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad ; \quad \nabla h(x^*) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda^* \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \lambda^* = \frac{1}{2}$$

Mirem condicions de 2n ordre:

$$\nabla^2 f(x^*) = 0 \quad ; \quad \nabla^2 h(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \underset{1/2}{\underset{\shortparallel}{\lambda^*}} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \succ 0 \Rightarrow x^* \text{ és mínim (punt òptim)}$$

3.- Caracterització de les direccions factibles

Considerem només les restriccions actives:

$h_i(x) = 0 \qquad i = 1, \dots, m$

$g_j(x) = 0 \qquad j \in A(x^*)$

Per tant, podem considerar que la regió factible està definida per

$\underbrace{c(x) = 0}_{\text{defineix una superfície } S \subseteq \mathbb{R}^n} \rightarrow c_i(x) = 0, \; i = 1, \dots, m' \quad (m' = m + |A(x^*)|)$

Ara hem de caracteritzar les direccions factibles a S en x^*:

1) Caracterització usant corbes factibles i diferenciables

Definició. Corba factible diferenciable dos cops
és a dir, $c(x(t)) = 0 \; \forall t \in [a,b]$
És una família de punts $x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \in S$, $a \leq t \leq b$ tal que
$\exists \; x'(t)$ i $x''(t)$. Si la corba passa per x^*, aleshores $\exists t^* \mid x(t^*) = x^*$.

Amb $x(t)$ podem obtenir direccions factibles. En el límit la direcció factible d és tangent a S en el punt x^*: $x'(t^*)$

Definició. Pla tangent en x^*.
El conjunt de direccions factibles equival al conjunt de derivades en x^* de totes les possibles corbes factibles i diferenciables que passen per x^*. Això s'anomena pla tangent T en x^*, i és un subespai de $\mathbb{R}^n$.

Definició. Subespai de direccions M.

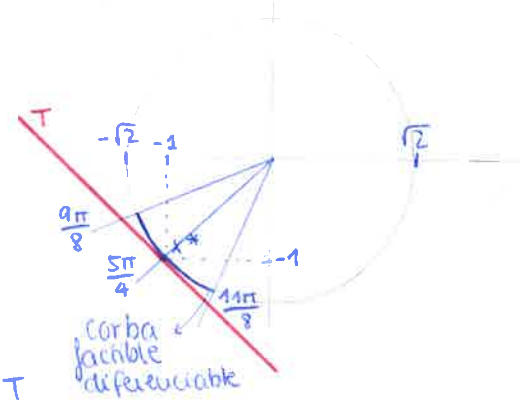
$M = \{ d \in \mathbb{R}^n \mid \nabla f(x^*)^T \cdot d = 0 \} = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i = 1, \dots, m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \end{array} \right\}$

$\downarrow$

És el conjunt de direccions ortogonals als gradients de les restriccions actives en el punt òptim.

Exemple.

$$\left.\begin{array}{l}\min\ x_1+x_2 \\ \text{s.a.:}\ x_1^2+x_2^2-2=0\end{array}\right\} \qquad x^*=\begin{pmatrix}-1\\-1\end{pmatrix}$$



La corba factible diferenciable és:

$$x(t)=\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix}=\begin{pmatrix}\sqrt{2}\cos t\\ \sqrt{2}\sin t\end{pmatrix} \qquad \frac{9\pi}{8}\leq t\leq\frac{11\pi}{8}$$

A $t^*=\frac{5\pi}{4}$, $x(t^*)=x^*$ i $x'(t^*)=d\in T$

$$d=x'(t^*)=\begin{pmatrix}x_1'(t)\\x_2'(t)\end{pmatrix}_{t=\frac{5\pi}{4}}=\begin{pmatrix}-\sqrt{2}\sin t\\ \sqrt{2}\cos t\end{pmatrix}_{t=\frac{5\pi}{4}}=\begin{pmatrix}1\\-1\end{pmatrix} \rightarrow \text{direcció factible.}$$

d és de descens? NO, ja que $\nabla f(x^*)^T\cdot d=(1\ \ 1)\begin{pmatrix}1\\-1\end{pmatrix}=1-1=0\not<0$

2) Caracterització usant seqüències factibles.

Donat x^*, una seqüència de punts $\{x_k\}_{k=0}^{\infty}$ on $x_k\in\mathbb{R}$ és factible si:

i) $x_k\neq x^*\ \ \forall k$

ii) $\lim\limits_{k\to\infty} x_k=x^*$

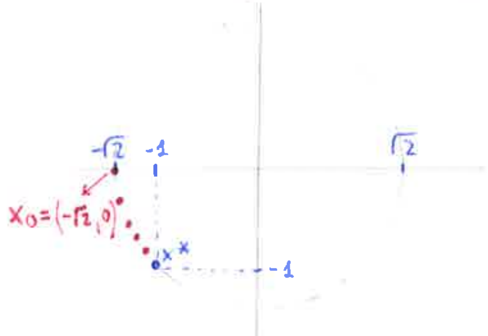
iii) $\exists K$ tq x_k és factible $\forall k\geq K$

Les direccions factibles són direccions límits de les direccions derivades de les seqüències factibles:

$$\lim_{k\to\infty}\frac{x_k-x^*}{\|x_k-x^*\|}=\lim_{k\to\infty}\frac{d_k}{\|d_k\|}=d$$

Exemple.

$$\left.\begin{array}{l}\min\ x_1+x_2 \\ \text{s.a.:}\ x_1^2+x_2^2-2=0\end{array}\right\} \qquad x^*=\begin{pmatrix}-1\\-1\end{pmatrix}$$



La seqüència factible diferenciable és:

$$x_k=\begin{pmatrix}-\sqrt{2-\left(1-\frac{1}{k+1}\right)^2}\\ -\left(1-\frac{1}{k+1}\right)\end{pmatrix}$$

i) $x_k\neq x^*\ \ \forall k$ $\underline{\text{ok}}$

ii) $\lim\limits_{k\to\infty} x_k=\begin{pmatrix}-1\\-1\end{pmatrix}=x^*$ $\underline{\text{ok}}$

iii) x_k és factible $\forall k\geq 0$: $\left(-\sqrt{2-\left(1-\frac{1}{k+1}\right)^2}\right)^2+\left(-\left(1-\frac{1}{k+1}\right)\right)^2=$

$$=2-\left(1-\frac{1}{k+1}\right)^2+\left(1-\frac{1}{k+1}\right)^2=2\ \ \underline{\text{ok}}$$

El punt $x_0 = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}$; la direcció límit és $\underbrace{d = \begin{pmatrix} -1 \\ 1 \end{pmatrix}}$

igual que quan hem usat la corba diferenciable factible.

Definició. Punt regular.

Un punt factible x és regular si els gradients de les restriccions actives $\nabla h_i(x)$ $i=1,\dots,m$ i $\nabla g_j(x)$ $j \in A(x)$ són linealment independents.

Treballar amb el pla tangent T o direccions límits de les seqüències factibles no és còmode. Cal una representació en funció de h i g. Per poder fer això cal imposar que el punt x^* sigui regular.

Teorema. Si x^* és punt regular en la superfície S definida per $c(x)=0$, llavors el pla tangent T de S en x^* és igual a M $(T=M)$.

Demostració

(tenim el problema $\min f(x)$ s.a.: $c(x)=0$)

$T \subseteq M$ | Donada $x(t)$ $a \leq t \leq b$, $x(t^*) = x^*$

no depèn de la regularitat de x^*

$x(t^*) \Rightarrow x'(t^*) \in M$

$c(\underbrace{x(t)}_{\text{factible}}) = 0 \quad \forall t \qquad \underbrace{\dfrac{d\,c(x(t))}{dt}}_{\text{constant}} = 0$

$$(\nabla c(\underbrace{x(t^*)}_{x^*}))^T \cdot x'(t^*) = \frac{dc(x(t))}{dt} = 0$$

$$\nabla c(x^*)^T \cdot x'(t^*) = 0 \Rightarrow x'(t^*) \in M \quad \text{OK}$$

necessitem la regularitat de x^*

$M \subseteq T$ | Sabem que x^* és regular.

Tenim $d \in M \Rightarrow \nabla c(x^*)^T \cdot d = 0$

$d \in T$? $\rightarrow$ hem de veure que $\exists\, x(t)$ $a \leq t \leq b$ tq $x(t^*) = x^*$ i $x'(t^*) = d$

↳ la demostració construeix $x(t)$.

considerem les equacions $c(x^* + t \cdot d + \nabla c(x^*) u) = 0$, $t \in \mathbb{R}$, $u \in \mathbb{R}^{m'}$, $\nabla c(x^*) = \begin{pmatrix} \nabla c_1 & \cdots & \nabla c_{m'} \end{pmatrix} \in \mathbb{R}^{n \times m'}$ $\Rightarrow$

$\Rightarrow$ tenim un sistema amb m' equacions i $m'+1$ variables.

Recordem el Teorema de la Funció Implícita:

$m \updownarrow$ $c(\underbrace{x}_{m}, \underbrace{y}_{n-m}) = 0$ — n variables, m equacions, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^{n-m}$

$\exists\, x^0, y^0 \mid c(x^0, y^0) = 0$. Si $\overbrace{\nabla_x c(x,y) \text{ és NO sing.}}^{\text{té inversa}}$

aleshores $\exists\, \Phi$ tq $x = \Phi(y)$ en un entorn de (x^0, y^0)

$x^0 = \Phi(y^0)$ i $c(\Phi(y^0), y^0) = 0$

Apliquem el Teorema de la Funció Implícita.

El nostre sistema $c(x^* + t\cdot d + \nabla c(x^*)u) = 0$ té solució si $\underbrace{t=0}_{y}$, $\underbrace{u=0}_{x}$ · ok

Calculem la derivada respecte u avaluat en $t=0$, $u=0$:

$$\nabla_u c(x^* + td + \nabla c(x^*)u) \underset{\substack{t=0\\u=0}}{=} \nabla c(x^* + \cancel{td} + \cancel{\nabla c(x^*)u})^T \cdot \nabla c(x^*) = \underbrace{\nabla c(x^*)^T \cdot \nabla c(x^*)}_{\in \mathbb{R}^{m' \times m'}}$$

$\nabla c(x^*)^T \cdot \nabla c(x^*)$ és singular? Sí ; x^* és regular, és a dir, els gradients de c són l.i. Si fem el producte també ens sortiran totes les columnes L.I, és a dir una matriu de rang complet ⇒ NO singular. OK

Per tant, pel T.F. Implícita $\exists\, u(t)$, $-\varepsilon \le t \le \varepsilon$ tq $\underset{(t=0)}{u(t)=0}$

La $\underbrace{\text{corba factible}}_{\text{És factible per construcció.}}$ serà $x(t) = x^* + td + \nabla c(x^*)\cdot u(t)$, on la única variable és la t.

Passa per l'òptim? Sí ; $x(0) = x^* + \cancel{0\cdot d} + \cancel{\nabla c(x^*)\cdot \overset{0}{u(0)}} = x^*$ OK

Podem garantir que $x'(\underset{0}{t^*}) = d$?

$$x'(t) = d + \nabla c(x^*)u'(t)$$

$$x'(0) = d + \underbrace{\nabla c(x^*)u'(0)}_{\text{hem de veure que això sigui zero.}} \Rightarrow u'(0) = 0 ? \text{ sí ;}$$

perquè és una constant

$$0 = \frac{dc(x(t))}{dt} = \nabla c(\overbrace{x(0)}^{x(t^*)=x^*})^T \cdot (d + \nabla c(x^*)u'(0)) =$$

$$= \underbrace{\nabla c(x^*)^T \cdot d}_{0} + \underbrace{\nabla c(x^*)^T \cdot \nabla c(x^*) \cdot u'(0)}_{} \Rightarrow u'(0) = 0 \Rightarrow$$

$\Rightarrow x'(t^*) = x'(0) = d$ OK

↳ matriu NO singular. vector hem vist que una solució és 0, i com que és SCD, 0 és la única solució OK

✱ El teorema anterior ens permet utilitzar els vectors $d \in M$ com a direccions factibles. La regularitat NO és una propietat de la superfície S, sinó de com es representa amb $c(x) = 0$. El pla tangent T depèn de S, el conjunt M depèn de $c(x) = 0$.

Exemple.

$\min\ x_1 + x_2 \rightarrow f(x)$

s.a.: $(x_1^2 + x_2^2 - 2)^2 = 0 \rightarrow h(x)$

$\nabla f(x) = \begin{pmatrix}1\\1\end{pmatrix}$; $\nabla h(x) = \begin{pmatrix}2(x_1^2+x_2^2-2)\cdot 2x_1\\ 2(x_1^2+x_2^2-2)\cdot 2x_2\end{pmatrix}$

$x^* = \begin{pmatrix}-1\\-1\end{pmatrix}$

$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$

$= \begin{pmatrix}1\\1\end{pmatrix} + \lambda^* \begin{pmatrix}0\\0\end{pmatrix} = 0$

$\nabla h(x^*) = \begin{pmatrix}0\\0\end{pmatrix}$ → NO és LI ⇒ x^* NO és regular

i $\nexists\, \lambda^*$ tq $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$.

PER TANT...? què vol dir això ???

4.- Condicions d'optimalitat

Teorema Condicions necessàries de 1r ordre (condicions KKT).

Donat $\min f(x)$ s.a.: $h(x)=0$, $g(x)\leq 0$ $\left.\right\}$ $f:\mathbb{R}^n\to\mathbb{R}\in\mathcal{C}^1$, $h:\mathbb{R}^n\to\mathbb{R}^m\in\mathcal{C}^1$, $g:\mathbb{R}^n\to\mathbb{R}^p\in\mathcal{C}^1$

i la seva lagrangiana $\mathcal{L}(x,\lambda,\mu) = f(x)+\lambda^T h(x)+\mu^T g(x) =$

$$= f(x)+\sum_{i=1}^{m}\lambda_i h_i(x)+\sum_{j=1}^{p}\mu_j g_j(x).$$

Suposem que x^* és un mínim local. Si x^* és punt regular, aleshores existeixen uns vectors $\lambda^*\in\mathbb{R}^m$ i $\mu^*\in\mathbb{R}^p$ tals que:

i) $h(x^*)=0$, $g(x^*)\leq 0$

ii) $\nabla_x\mathcal{L}(x^*,\lambda^*,\mu^*) = \nabla f(x^*)+\lambda^*\nabla h(x^*)+\mu^*\nabla g(x^*) = 0.$

iii) $\mu^*\geq 0$ i $\mu^{*T}\cdot g(x^*)=0$ (si $j\notin A(x^*)$ llavors $\mu_j^*=0$)

<u>Demostració</u>

i) <u>OK</u> (per a ser òptim ha de ser factible)

ii),iii) Si s'ha de garantir $\mu_j^* g_j(x^*)=0$ $j=1,\dots,p$ $\Rightarrow$

$\Rightarrow$ si $j\notin A(x^*)$ $(g_j(x^*)<0)$, aleshores $\mu_j^*=0$. només ens calen les μ de les restriccions actives!

Hem de provar que $\nabla f(x^*)+\sum_{i=1}^{m}\lambda_i^*\nabla h_i(x^*)+\underbrace{\sum_{j\in A(x^*)}\mu_j^*\nabla g_j(x^*)}=0$

simplifiquem la notació: $\nabla f(x^*)+\sum_{i=1}^{m'}\lambda_{c_i}^*\nabla c_i(x^*)=0$ $\quad(m'=m+|A(x^*)|)$

$$\left(c(x)=\begin{pmatrix} h_1(x)\\ \vdots \\ h_m(x)\\ g_j(x)\ \ j\in A(x^*)\end{pmatrix};\quad \lambda_c=\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_m\\ \mu_j\ \ j\in A(x^*)\end{pmatrix}\right)$$

Provarem que si $d\in M=\{d\mid \nabla c(x^*)^T\cdot d=0\}$, aleshores $\nabla f(x^*)^T\cdot d=0$.

Sabem que x^* és regular. Aleshores, pel teorema anterior, $T=M$.

Llavors $\exists\, x(t)$ $a\leq t\leq b$ tq $x(t^*)=x^*$ i $x'(t^*)=d$

$$0=\left.\frac{df(x(t))}{dt}\right|_{t=t^*}=\nabla f(x(t^*))^T\cdot x'(t^*) = \nabla f(x^*)^T\cdot d = 0 \quad \forall d\in M.\ \ \underline{\text{ok}}.$$

M, de fet, és el nucli de $\nabla c(x^*)$ i té dimensió $n-m'$

Imatge de $\nabla c(x^*)$ és $I=\{v\in\mathbb{R}^n\mid v=\sum_{i=1}^{m'}\lambda_i\nabla c_i(x^*)\}$ de dimensió m' (perquè x^* regular).

Sabem que $\left.\begin{array}{l} I \perp M \\ \nabla f(x^*) \perp M \end{array}\right\} \Rightarrow \nabla f(x^*) \in I \Rightarrow \exists \lambda_{ci} \quad i = 1,\dots,m'$

tq $\quad 0 = \nabla f(x^*) + \lambda_{ci}^* \nabla c(x^*) \quad i = 1,\dots,m'$

Ara només ens falta verificar que $\mu_j^* \geq 0 \quad j = 1,\dots,p$ i que $\mu_j^* = 0 \quad j \notin A(x^*)$.

↳ reducció a l'absurd: suposem que $\exists k$ tq $\mu_k^* < 0$ (veurem que x^* NO és l'òptim)

Considerem

$$S' = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} h_i(x) = 0 & i = 1,\dots,m \\ g_j(x) = 0 & j \in A(x^*) \setminus \{k\} \end{array} \right\}$$

$$M' = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i = 1,\dots,m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \setminus \{k\} \end{array} \right\}$$

Sabem que $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j \in A(x^*) \setminus \{k\}} \mu_j^* \nabla g_j(x^*) + \mu_k^* \nabla g_k(x^*) = 0$

Considerem $d \in M'$, i suposem que per a la restricció k $\nabla g_k(x^*)^T \cdot d < 0 \rightarrow$ OK

(d factible per a S') $-d$ també.

En el cas que $\nabla g_k(x^*)^T \cdot d > 0$, considerarem $-d$.

↳ sempre $\exists$ una direcció factible i de descens

perquè ens està tirant cap a l'interior de la restricció, per tant és factible $\Rightarrow$ d és de descens $\Rightarrow$ x^* NO és l'òptim

Així doncs, $d \in M'$ factible i x^* regular $\Rightarrow T' = M'$ i $\exists x(t) \quad a \leq t \leq b$

tq $x(t^*) = x^*$ i $x'(t^*) = d$

$$\left.\frac{df(x(t))}{dt}\right|_{t=t^*} = \frac{df(x^*)}{dt^*} = \nabla f(x^*)^T \cdot x'(t^*) = \nabla f(x^*)^T \cdot d < 0 \quad \text{OK}$$

Teorema. Condicions necessàries de 2n ordre.

Donat $\left.\begin{array}{l} \min f(x) \\ \text{s.a.: } h(x) = 0 \\ \qquad g(x) \leq 0 \end{array}\right\}$ on $f, h, g \in \mathcal{C}^2$. Si x^* és mínim local i és un punt regular, aleshores es satisfan les condicions KKT de 1r ordre i a més

$$\nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x) + \sum_{j=1}^{p} \mu_j^* \nabla^2 g_j(x^*) \succcurlyeq 0$$

$\forall d \in M = \left\{ d \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i = 1,\dots,m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \end{array} \right\}$ és a dir, $d^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*) \cdot d \geq 0 \quad \forall d \in M$.

Teorema. Condicions suficients de 2n ordre.

Donat $\left.\begin{array}{l}\min f(x) \\ \text{s.a.: } h(x)=0 \\ \quad g(x)\leq 0\end{array}\right\}$ on $f, h, g \in \mathcal{C}^2$. Són condicions suficients per a que

x^* sigui mínim local estricte que:

1) es satisfacin les condicions KKT de 1r ordre en x^* (no cal que x^* sigui regular)

2) $\nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla^2 g_j(x^*) \succ 0$

$\forall d \in M' = \left\{ d \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i=1,\dots,m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \cap \{j \mid \mu_j^* > 0\} \end{array} \right\}$; és a dir,

$d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)\, d > 0 \;\; \forall d \in M'$.

En resum:

Donat el problema $\left.\begin{array}{l}\min f(x) \\ \text{s.a.: } h(x)=0 \\ \quad g(x)\leq 0\end{array}\right\}$

i la seva lagrangiana $\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$

- condicions necessàries: si x^* punt regular és òptim local, llavors han d'existir $\lambda^* \in \mathbb{R}^m$ i $\mu^* \in \mathbb{R}^p$ tals que

  i) $h(x^*) = 0$, $g(x^*) \leq 0$

  ii) $\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0$

  iii) $\mu^* \geq 0$ i $\mu^{*T} \cdot g(x^*) = 0$ (si $g_j(x^*)$ és inactiva llavors $\mu_j^* = 0$)

  condicions de 2n ordre

  iv) $d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)\, d \geq 0 \quad \forall d \in M = \left\{ d \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i=1,\dots,m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \end{array} \right\}$

- condicions suficients: el punt x^* és òptim local si satisfà:

  condicions 1r ordre KKT → NO CAL IMPOSAR QUE SIGUI REGULAR.

  i) $h(x^*) = 0$, $g(x^*) \leq 0$

  ii) $\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0$

  iii) $\mu^* \geq 0$ i $\mu^{*T} g(x^*) = 0$ (si $g_j(x^*) < 0 \Rightarrow \mu_j^* = 0$)

  condicions 2n ordre

  iv) $d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)\, d > 0 \quad \forall d \in M' = \left\{ d \;\middle|\; \begin{array}{ll} \nabla h_i(x^*)^T \cdot d = 0 & i=1,\dots,m \\ \nabla g_j(x^*)^T \cdot d = 0 & j \in A(x^*) \cap \{j \mid \mu_j^* > 0\} \end{array} \right\}$

Exemple. Donat un segment de longitud a, volem dividir-lo en dues parts, de manera que la suma de les àrees dels quadrats de costat cada una de les parts sigui mínima.

$$\left.\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.a.:} & x_1 + x_2 = a \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}\right\}$$

(entenem que x_1 i x_2 serà la longitud de cada part)

$$\left.\begin{array}{ll} \min & x_1^2 + x_2^2 \to f(x) \\ \text{s.a.:} & x_1 + x_2 - a = 0 \to h(x) \\ & -x_1 \leq 0 \to g_1(x) \\ & -x_2 \leq 0 \to g_2(x) \end{array}\right\}$$

Lagrangiana:

$$\mathcal{L}(x, \lambda, \mu) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - a) + \mu_1(-x_1) + \mu_2(-x_2).$$

- Condicions necessàries de 1r ordre:

i) $x_1 + x_2 - a = 0$; $-x_1 \leq 0$; $-x_2 \leq 0$

ii) $\nabla_x \mathcal{L}(x, \lambda, \mu) = 0$

$\hookrightarrow$ $\left.\begin{array}{l} 2x_1 + \lambda - \mu_1 = 0 \\ 2x_2 + \lambda - \mu_2 = 0 \end{array}\right\}$

iii) $\mu_i \geq 0$; $\mu_i = 0$ si $g_i(x) = 0$ $i = 1,2$.

- Hem d'analitzar els 4 casos segons $g(x)$.

a) $x_1 = 0$, $x_2 = 0$

$\hookrightarrow$ No pot ser, infactibilitat $x_1 + x_2 = a$ $(a > 0)$

b) $x_1 > 0$, $x_2 = 0$

$\hookrightarrow$ Aleshores $x_1 = a$ i $\mu_1 = 0$.

Si resolem el sistema, $\left.\begin{array}{l} 2a + \lambda = 0 \\ \lambda - \mu_2 = 0 \end{array}\right\}$ $\lambda = \underbrace{\mu_2 = -2a}_{\text{No pot ser, ja que } \mu_2 < 0.}$

c) $x_2 > 0$, $x_1 = 0$

$\hookrightarrow$ No pot ser per simetria amb el cas b)

d) $x_1 > 0$, $x_2 > 0$

$\hookrightarrow$ Aleshores $\mu_1 = \mu_2 = 0$.

Si resolem el sistema, $\left.\begin{array}{l} 2x_1 + \lambda = 0 \\ 2x_2 + \lambda = 0 \\ x_1 + x_2 = a \end{array}\right\}$ $\begin{array}{l} \lambda = -2x_1 = -2x_2 \\ x_1 = x_2 \\ 2x_1 = a \Rightarrow x_1 = a/2 = x_2 \end{array}$

Per tant $\lambda = -2x_1 = -a$

$\hookrightarrow$ Es compleixen totes les condicions, tenim un candidat a òptim.

- Condicions suficients de 2n ordre:

$$\nabla^2_{xx} \mathcal{L}(x,\lambda,\mu) = \nabla^2 f(x) + \lambda \nabla^2 h(x) + \mu_1 \nabla^2 g_1(x) + \mu_2 \nabla^2 g_2(x).$$

$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ $\quad \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

$\nabla h(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\quad \nabla^2 h(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\nabla g_1(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ $\quad \nabla^2 g_1(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $\quad \nabla^2 g_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Per tant, $\nabla^2_{xx} \mathcal{L}(x,\lambda,\mu) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \succ 0 \quad \forall x \quad \Rightarrow$

$\Rightarrow$ la hessiana de la lagrangiana és definida $\forall d$, no només $d \in M$.

Per tant, el punt obtingut és mínim local (i com que el problema és convex, també global). $x^* = \begin{pmatrix} a/2 \\ a/2 \end{pmatrix}$.

Exemple. Volem construir un cilindre metàl·lic de volum màxim, de forma que la seva superfície sigui 6π.

↳ variables: radi de la base (r), alçada del cilindre (h).

$$\left.\begin{array}{ll} \max & \pi \cdot r^2 \cdot h \\ \text{s.a.:} & 2\pi r \cdot h + 2\pi r^2 = 6\pi \\ & r \geq 0 \\ & h \geq 0 \end{array}\right\}$$

$$\left.\begin{array}{lll} \min & -\pi \cdot r^2 \cdot h & \rightarrow \cancel{f(x)}\ f(r,h) \\ \text{s.a.:} & 2\pi r h + 2\pi r^2 - 6\pi = 0 & \rightarrow \cancel{h(x)}\ h(r,h) \\ & -r \leq 0 & \rightarrow \cancel{g_1(x)}\ g_1(r,h) \\ & -h \leq 0 & \rightarrow \cancel{g_2(x)}\ g_2(r,h) \end{array}\right\}$$

Lagrangiana:

$$\mathcal{L}(x,\lambda,\mu) = f(x) + \lambda h(x) + \mu_1 g_1(x) + \mu_2 g_2(x).$$

$x = r, h$

- Condicions necessàries de 1r ordre:

i) S'han de complir les restriccions (factibilitat)

ii) $\nabla_{r,h} \mathcal{L}(r,h,\lambda,\mu) = 0$

↳ $\nabla_r \mathcal{L}(r,h,\lambda,\mu) = -2\pi r h + 2\pi h\lambda + 4\pi r\lambda - \mu_1 = 0$

↳ $\nabla_h \mathcal{L}(r,h,\lambda,\mu) = -\pi r^2 + 2\pi r\lambda - \mu_2 = 0$

iii) $\mu_1 \cdot g_1(r,h) = 0, \quad \mu_2 \cdot g_2(r,h) = 0.$

g_1 i g_2 han de ser inactives ($r, h > 0$) ja que sinó la solució és infactible i/o NO òptima. Així doncs, $\mu_1, \mu_2 = 0$.

Ara resolem el sistema:

$$\left.\begin{array}{l} 1)\ -2\pi r h + 2\pi h \lambda + 4\pi r \lambda = 0 \\ 2)\ -\pi r^2 + 2\pi r \lambda = 0 \\ 3)\ 2\pi r h + 2\pi r^2 - 6\pi = 0 \end{array}\right\}$$

2) $2\cancel{\pi}\cancel{r}\lambda = \cancel{\pi} r^{\cancel{2}}$

$2\lambda = r \Rightarrow \lambda = \frac{r}{2}$

1) $-2\pi r h + \cancel{2}\pi h \frac{r}{\cancel{2}} + 4\pi r \frac{r}{2} = 0$

$-\pi r h + 2\pi r^2 = 0$

$h = \frac{2\cancel{\pi} r^2}{\cancel{\pi}\cancel{r}} \Rightarrow h = 2r$

3) $2\pi r(2r) + 2\pi r^2 - 6\pi = 0$

$4\pi r^2 + 2\pi r^2 - 6\pi = 0$

$6\pi r^2 - 6\pi = 0 \Rightarrow r^2 = \frac{6\pi}{6\pi} = 1 \Rightarrow \boxed{r = 1} \quad r \cancel{=} -1$

Per tant, $h = 2r \Rightarrow \boxed{h = 2}$ i $\lambda = r/2 \Rightarrow \boxed{\lambda = 1/2}$

↳ Aquest punt satisfà les condicions de 1r ordre, i és candidat a òptim

- Comprovem les condicions de 2n ordre: $\left(d^T \nabla^2_{r,h} \mathcal{L}(x,\lambda,\mu) \cdot d > 0 \ \forall d \mid \nabla c(x)^T \cdot d = 0 \right)$

$\nabla f(r,h) = \begin{pmatrix} -2\pi r h \\ -\pi r^2 \end{pmatrix} \qquad \nabla^2 f(r,h) = \begin{pmatrix} -2\pi h & -2\pi r \\ -2\pi r & 0 \end{pmatrix}$

$\nabla h(r,h) = \begin{pmatrix} 2\pi h + 4\pi r \\ 2\pi r \end{pmatrix} \qquad \nabla^2 h(r,h) = \begin{pmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{pmatrix}$

$\nabla g_1(r,h) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \qquad \nabla^2 g_1(r,h) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\nabla g_2(r,h) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \qquad \nabla^2 g_2(r,h) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$$\nabla^2_{r,h} \mathcal{L}(r,h,\lambda,\mu)\Big|_{\substack{r=1\\h=2\\\lambda=1/2}} = \left. \begin{pmatrix} -2\pi h & -2\pi r \\ -2\pi r & 0 \end{pmatrix} + \lambda \begin{pmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{pmatrix} \right|_{\substack{r=1\\h=2\\\lambda=1/2}} =$$

$$= \begin{pmatrix} -4\pi & -2\pi \\ -2\pi & 0 \end{pmatrix} + 1/2 \begin{pmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{pmatrix} = \begin{pmatrix} -2\pi & -\pi \\ -\pi & 0 \end{pmatrix}$$

Vectors d tals que $\nabla c(x)^T \cdot d = 0$:

$$\nabla c(x) = \nabla_{r,h} h(r,h)\Big|_{\substack{r=1\\h=2}} = \begin{pmatrix} 2\pi h + 4\pi r \\ 2\pi r \end{pmatrix}\Bigg|_{\substack{r=1\\h=2}} = \begin{pmatrix} 8\pi \\ 2\pi \end{pmatrix}$$

$$\begin{pmatrix} 8\pi & 2\pi \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0 \Rightarrow 8\pi d_1 + 2\pi d_2 = 0 \Rightarrow d_2 = -4 d_1$$

$$d^T \cdot \nabla^2_{r,h} \mathcal{L}(x,\lambda,\mu) \cdot d = \begin{pmatrix} d_1 & -4d_1 \end{pmatrix} \begin{pmatrix} -2\pi & -\pi \\ -\pi & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ -4d_1 \end{pmatrix} = 6\pi d_1^2 > 0 \ \forall d \neq 0$$

Per tant ens garanteix una condició suficient de 2n ordre. $x^* = \begin{pmatrix} r = 1 \\ h = 2 \end{pmatrix}$ és mínim local. ok

Teorema. Anàlisi de sensibilitat

Donada la família de problemes $\left.\begin{array}{l}\min f(x)\\ \text{s.a.: } h(x) = \varepsilon \\ \quad g(x) \leq \delta\end{array}\right\}$ on $f, h, g \in \mathcal{C}^2$.

Suposem que per a $\varepsilon = \delta = 0$ hi ha una solució x^*, punt regular, i λ^*, μ^* que satisfan les condicions d'optimalitat KKT de 1r ordre i les suficients de 2n ordre. Suposem també que si $j \in A(x^*)$ llavors $\mu_j^* > 0$.

Aleshores, $\forall (\varepsilon, \delta)$ en un entorn de $(0,0)$ hi ha una solució $x^*(\varepsilon, \delta)$ tal que $x^*(0,0) = x^*$. A més es verifica
$$\nabla_\varepsilon f(x(\varepsilon,\delta))\Big|_{(0,0)} = -\lambda^*$$
$$\nabla_\delta f(x(\varepsilon,\delta))\Big|_{(0,0)} = -\mu^*$$

Exemple. Si ens fixem en l'últim exemple,

No ho entenc

$\left.\begin{array}{l}\min f(r,h) = -\pi r^2 h\\ \text{s.a.: } c(r,h) = 2\pi r h + 2\pi r^2 - 6\pi = 0\\ \quad r \geq 0\\ \quad h \geq 0\end{array}\right\}$ solució òptima: $r = 1$, $h = 2$, $\lambda = 1/2$ $\quad f(r,h) = -2\pi$.

Considerem $c(r,h) = 2\pi r h + 2\pi r^2 - 6\pi = 0'01$.

El nou valor òptim és $f(r(\varepsilon), h(\varepsilon)) = -2\pi - 0'0050007$, i
$$\Delta f = f(r(\varepsilon), h(\varepsilon)) - f(r,h) = -0'0050007$$
Observem que $\Delta f \approx \varepsilon(-\lambda) = 0'01(-1/2) = -0'005$

5.- Problemes convexos

Definició. Problema estrictament convex

Donat $\left.\begin{array}{l}\min f(x)\\ \text{s.a.: } h(x) = 0\\ \quad g(x) \leq 0\end{array}\right\}$ on $\left\{\begin{array}{l} f \text{ estrictament convexa}\\ h(x) = Ax - b \text{ funció afí}\\ g_j \text{ convexa } j = 1, \dots, p\end{array}\right.$

Aleshores $\nabla^2 f(x) \succ 0$, $\nabla^2 g_j(x) \succeq 0$ $j = 1, \dots, p$, $\nabla^2 h(x) = 0$

Per tant, $\nabla^2_{xx} \mathcal{L}(x, \lambda, \mu) = \underbrace{\nabla^2 f(x)}_{\succ 0} + \underbrace{\cancel{\sum_{i=1}^m \lambda_i \nabla^2 h_i(x)}}_{0} + \underbrace{\sum_{i=1}^p \overset{\geq 0}{\mu_j} \overset{\succeq 0}{\nabla^2 g_j(x)}}_{\succeq 0} \succ 0 \Rightarrow$

$\Rightarrow$ la hessiana de la lagrangiana és def. positiva $\Rightarrow$ verifica les condicions suficients.

És a dir, qualsevol punt que satisfaci KKT de 1r ordre, també satisfà les de 2n ordre i serà mínim global estricte.

Definició. Problema ("només") convex.

Sense imposar convexitat estricta també es garanteix que en un problema convex amb les condicions de 1r ordre n'hi ha prou per a obtenir el mínim global.

$\nabla^2_{xx} \mathcal{L}(x, \lambda, \mu) \succeq 0$

Teorema.

Donat $\left.\begin{array}{l}\min f(x)\\ \text{s.a.: } h(x)=0\\ g(x)\leq 0\end{array}\right\}$ $f,h,g \in \mathcal{C}^1$, f convexa, g_j convexes, $h(x)=Ax-b$ funció afí.

Si es verifiquen les condicions KKT de 1r ordre a x^*, llavors x^* és mínim global.

Demostració

Considerem $\underbrace{h(x)=0}_{\text{funció afí}}$ $\left\{\begin{array}{l} h(x)\leq 0\\ h(x)\geq 0\end{array}\right. \to \begin{array}{ll} h(x)\leq 0 & \text{convexa}\\ -h(x)\leq 0 & \text{convexa}\end{array}$ $\overset{\text{OK}}{\Rightarrow}$ no ens cal posar restriccions d'igualtat perquè venen englobades en el cas de desigualtat.

Prenem x^*, que satisfà les condicions KKT de 1r ordre $\Rightarrow$

$$\Rightarrow \left\{\begin{array}{l}\nabla f(x^*)+\nabla g(x^*)\mu^*=0\\ \mu^*\geq 0\\ g(x^*)^T\cdot\mu^*=0\end{array}\right.$$

Considerem un punt qualsevol factible $\bar{x}$.

$$\underset{\text{OK}}{f(\bar{x})\geq} f(\bar{x})+\underbrace{\underbrace{\mu^{*T}}_{\geq 0}\cdot\underbrace{g(\bar{x})}_{\leq 0}}_{\leq 0}$$

Recordem la definició de funció convexa:

$$f(x)\geq f(x')+\nabla f(x')(x-x')\quad \forall x,x'$$

Com que f,g convexes, $f(\bar{x})\geq f(x^*)+\nabla f(x^*)(\bar{x}-x^*)+\sum_{j=1}^{p}\mu_j^*(g_j(x^*)+\nabla g_j(x^*)(\bar{x}-x^*)) =$

$$= f(x^*)+\underbrace{\left(\nabla f(x^*)+\sum_{j=1}^{p}\mu_j^*\nabla g_j(x^*)\right)(\bar{x}-x^*)}_{=0 \,\leftarrow\, \text{KKT (i)}}+\underbrace{\sum_{j=1}^{p}\mu_j^* g_j(x^*)}_{=0\,\leftarrow\,\text{KKT (iii)}} = f(x^*)$$

Per tant, $f(\bar{x})\geq f(x^*)$ $\forall \bar{x}$ factible $\Rightarrow$ x^* és l'òptim global. OK

Problemes lineals i quadràtics

- Condicions d'optimalitat d'un PL $\left(\begin{array}{ll}\min c^Tx & \\ \text{s.a.: } Ax=b & \to \lambda\\ -x\leq 0 & \to \mu\end{array}\right)$

i) $Ax=b$, $x\geq 0$ } factibilitat primal

ii) $A^T\lambda+\mu=c$, $\mu\geq 0$ } factibilitat dual

iii) $\mu^T\cdot x=0$ } complementarietat

Lagrangiana: $\mathcal{L}(x,\lambda,\mu)=c^Tx+\lambda^T(b-Ax)-\mu^Tx$

$\hookrightarrow \nabla_x\mathcal{L}(x,\lambda,\mu)=c^T+\lambda^TA-\mu^T = c+A^T\lambda+\mu \underset{\text{(KKT)}}{=} 0$

Algorisme del símplex:

1) $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} \begin{matrix} \updownarrow n \\ \updownarrow n-m \end{matrix}$ $\quad x_B > 0$ (no hi ha degeneració), $\quad x_N = 0$

complementarietat

$\mu = \begin{bmatrix} \mu_B \\ \mu_N \end{bmatrix} \begin{matrix} \updownarrow n \\ \updownarrow n-m \end{matrix}$ $\quad \mu_B = 0$ → perquè $x^T \mu = 0$, $\quad \mu_N \geq 0$

↳ $\underbrace{x_B^T \mu_B}_{\downarrow} + \cancel{\underbrace{x_N^T}_{0} \mu_N} = 0$

$x_B > 0 \Rightarrow \mu_B = 0$. OK

2) $Ax = b$

↳ $Ax = [B \;\; A_N] \begin{bmatrix} x_B \\ x_N \end{bmatrix} = B x_B + \cancel{A_N \overset{0}{x_N}} = B x_B = b \Rightarrow x_B = B^{-1} \cdot b$

3) $A^T \cdot \lambda + \mu = c$

↳ $\begin{bmatrix} B^T \\ A_N^T \end{bmatrix} \lambda + \begin{bmatrix} \overset{0}{\mu_B} \\ \mu_N \end{bmatrix} = \begin{bmatrix} c_B \\ c_N \end{bmatrix} \Rightarrow \begin{cases} B^T \cdot \lambda = c_B \Rightarrow \overbrace{\lambda = B^{-T} \cdot c_B}^{\text{manera de calcular les variables duals}} \\ A_N^T \cdot \lambda + \mu_N = c_N \end{cases}$

ho reescrivim: $\mu_N = c_N - A_N^T \cdot \lambda \Rightarrow \mu_N = c_N - A_N^T (B^{-T} \cdot c_B) \Rightarrow$

Equació dels costos reduïts de les variables no bàsiques

$\Rightarrow$ si ho calculem així estem garantint que es compleixi $A^T \lambda + \mu = c$ (KKT)

El que ens falta veure és $\mu \geq 0$:

$\mu = \begin{bmatrix} \mu_B \to 0 \\ \mu_N \to ? \end{bmatrix} \Rightarrow$ per a que sigui factible (dual) han de ser ≥ 0.

Per tant, mentre anem iterant pel símplex primal i no estiguem a l'òptim, es compliran totes les condicions menys $\mu \geq 0$ (per això ens aturem quan els costos reduïts són $\geq 0 \Rightarrow$ hem arribat a l'òptim $\begin{cases} \text{factibilitat primal} \\ \text{factibilitat dual} \end{cases}$)

- PL

$$\begin{array}{ll} \min & c^T x \\ \text{s.a} & Ax = b \quad [\lambda] \\ & -x \leq 0 \quad [\mu] \end{array}$$

- Lagrangiana

$$L(x, \lambda, \mu) = c^T x + \lambda^T (b - Ax) - \mu^T x$$

- Condicions KKT

$$\begin{array}{l} Ax = b, \quad x \geq 0 \\ c - A^T \lambda - \mu = 0 \\ \mu^T x = 0, \quad \mu \geq 0 \end{array}$$

- Reescrivim KKT:

$$\begin{array}{ll} Ax = b, \quad x \geq 0 & \text{Factibilitat primal} \\ A^T \lambda + \mu = c, \quad \mu \geq 0 & \text{Factibilitat dual} \\ \mu^T x = 0 & \text{Complementarietat: } c^T x = b^T \lambda \end{array}$$

- Condicions KKT:

$$\begin{array}{l} Ax = b \\ A^T \lambda + \mu = c \\ \mu^T x = 0 \\ x \geq 0, \mu \geq 0 \end{array}$$

- Al símplex (primal) μ són els costos reduïts i λ les variables duals
- El símplex busca una partició entre bàsiques i no bàsiques, que verifica les condicions KKT:

$$\begin{array}{l} x = [x_B \;\; x_N]^T \quad x_N = 0 \quad x_B > 0 \\ A = [B \;\; N], \quad Ax = Bx_B = b \\ \mu_B = 0 \Rightarrow B^T \lambda = c_B \\ \mu_N = c_N - N^T \lambda [\text{ definició cost reduït símplex }] \\ \mu_N \geq 0 \Rightarrow \text{ Condició optimalitat símplex} \end{array}$$

- La complementarietat $\mu^T x = 0$ garantida a tota iteració del símplex: $\mu_B^T x_B = 0, \mu_N^T x_N = 0$
- Es viola (al símplex primal) la condició $\mu_N \geq 0$. Es va iterant fins aconseguir-la.