Optimització Estàtica

Xavier Martinez-Giralt
Universitat Autònoma de Barcelona

Optimització Estàtica

Índex

Índex de figures

Índex de taules

Capítol 1

Introducció.

Els problemes d'optimització impregnen el mon modern doncs apareixen tant a les ciències, i a les ciències socials com a la ingenieria. En particular, la utilitat de les tècniques d'optimització en economia apareix en la resolució dels problemes d'assignació de recursos. Encara que, com és natural, ens podem plantejar problemes estàtics i dinàmics, en aquest semestre només abordarem els primers. Així, estudiarem amb detall l'optimització estàtica que inclou la programació clàssica, la programació no lineal i la lineal. Un aspecte que complerta aquesta visió com és una introducció a la teoria de jocs la deixarem de banda per mes endavant. Les tècniques que estudiarem tenen aplicacions a problemes de la teoria del consum, teoria de l'empresa, l'equilibri general i l'economia del benestar.

Com ja s'ha mencionat una primera distinció important en els problemes d'optimització resideix entre els problemes estàtics i dinàmics. Un altra distinció es refereix al tipus de restriccions a que estan sotmesos els problemes d'optimització. Podem distinguir entre problemes sense restriccions o sotmesos a restricions d'igualtat del tipus

$$8x_1 + 2x_2 = 5$$

i problemes sotmesos a restriccions de desigualtat com per exemple

$$8x_1 + 2x_2 5$$

Una tercera distinció important es refereix al número d'agents que prenen decisions. Podem pensar en problemes Amb un únic agent decisor, o en problemes amb varios agents decisors.

Un tipus diferent de problemes que tampoc tractarem aquí són els problemes estocàstics. Nosaltres ens concentrarem en alguns problemes determinístics. Podem doncs resumir els diferents tipus de problemes determinístics d'optimització en la taula 1.

Restriccions;# Agents/Temps	Problemes Estàtics	Problemes Dinàmics
No Restriccions o d'igualtat Un agent decisor	Programació Clàssica	Càlcul de Variacions
Restriccions de desigualtat Un agent decisor	Programació No Lineal i Lineal	Teoria de Jocs
Dos o més agents decisors	Prog. Dinàmica i Principi del Màxim	Jocs Diferencials

Taula 1: Tipologia de problemes determinístics d'optimització.

El problema econòmic.

El problema bàsic en economia és l'assignació de recursos escasos entre objectius alternatius. Donada la manca de recursos, hem de fer una elecció, i aquesta elecció ha de ser racional, és a dir, hem de prendre decisions que permetin assolir objectius dins d'aquesta manca de recursos. Per exemple la distribució de renda entre consum i estalvi per els individus; el repartiment de la despesa en consum entre els diferents bens a l'abast, etc.

Aquest problema d'utilització eficient (racional) de recursos escasos no es més que un problema d'optimització definit com l'elecció de valors de certes variables per tal de maximitzar una funció subjecte a restriccions.

Les variables del problema econòmic són *instruments* que resumeixen la decisió d'una assignació particular; la funció a optimitzar és la *funció objectiu* que resumeix els objectius alternatius; i les *restriccions* resumeixen la manca de recursos; El *conjunt d'oportunitats* és el conjunt d'instruments que satisfan totes les restriccions. En termes matemàtics doncs, el problema econòmic no és més que la selecció dels instruments dins del conjunt d'oportunitats, que permetin optimitzar la funció objectiu.

Capítol 2

El Problema de la Programació Matemàtica.

El problema econòmic estàtic consisteix en assignar recursos escasos entre objectius alternatius (i competidors) en un moment particular del temps. Matemàticament, el problema consisteix en determinar el valor de certes variables subjectes a un conjunt de restriccions pre-establert sobre els posibles valors que poden prendre les variables, de manera que s'optimitzi una certa funció. Presentat en aquests termes, el problema econòmic estàtic també s'anomena *el problema de la programació matemàtica*.

2.1 Definició Formal del Problema.

L'enunciat formal del problema de la programació matemàtica compren els *instruments*, el *conjunt d'oportunitats*, i la *funció objectiu*.

El problema tracta d'escollir els valors per n variables $x_1, x_2, \ldots x_n$, denominats *instruments*. Els instruments els representem per un vector columna:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \big(x_1, x_2, \ldots, x_n\big)'$$

que denominem el vector d'instruments, que és un vector en l'espai Euclidi E^n.

El vector d'instruments $\mathbf{x}$ és *factible* si satisfà totes les restriccions el problema, i el conjunt de tots els vectors d'instruments factibles es denomina el *conjunt d'oportunitats*, X, un subconjunt de E^n. Donat que el problema que tenim que resoldre consisteix en escollir un vector d'instruments del conjunt d'oportunitats, en qualsevol problema que no sigui trivial, el conjunt d'oportunitats no serà buit (i.e. el conjunt de restriccions no serà inconsistent) i contindrà al menys dos elements.

La *funció objectiu* és una funció amb valor real dels instruments:

$$F = f(\mathbf{x}) = F(x_1, x_2, \ldots, x_n)$$

El problema general de la programació matemàtica el podem formular aleshores com la selecció d'un vector d'instruments dins del conjunt d'oportunitats per tal de maximitzar el valor de la funció objectiu:

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad subjecte \;\; a \quad \mathbf{x} \in X$$

Casos particulars especialment importants d'aquest problema general són la programació clàssica, la programació no lineal i la programació lineal.

- En la *programació clàssica* les restriccions són del tipus d'igualtat, i consisteixen en m igualtats:

$$\begin{pmatrix} g_1(\mathbf{x}) = g_1(x_1, x_2, \ldots, x_n) = b_1 \\ g_2(\mathbf{x}) = g_2(x_1, x_2, \ldots, x_n) = b_2 \\ \vdots \\ g_m(\mathbf{x}) = g_m(x_1, x_2, \ldots, x_n) = b_m \end{pmatrix}$$

  on les funcions $g_i(\mathbf{x}),\ i = 1, 2, \ldots, m,$ són funcions contínuament diferenciables dels instruments denominades *restriccions*, i els paràmetres $b_i,\ i = 1, 2, \ldots, m,$ són m números reals denominats *constants*. En notació vectorial podem escriure les restriccions com

$$\mathbf{g}(\mathbf{x}) = \mathbf{b}$$

  on $\mathbf{g}(\mathbf{x})$ i $\mathbf{b}$ són els vectors columna m-dimensionals:

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_1(x_1, x_2, \ldots, x_n) \\ g_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

  Així doncs, el problema de la programació clàssica consisteix en maximitzar una certa funció subjecte a restriccions d'igualtat:

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad subjecte \;\; a \quad \mathbf{g}(\mathbf{x}) = \mathbf{b}$$

- En la *programació no lineal* les restriccions són de dos tipus: restriccions de *no negativitat* del tipus

  $$x_1 0, x_2 0, \ldots, x_n 0$$

  i restriccions de *desigualtat*:

  $$\begin{pmatrix} g_1(\mathbf{x}) = g_1(x_1, x_2, \ldots, x_n) b_1 \\ g_2(\mathbf{x}) = g_2(x_1, x_2, \ldots, x_n) b_2 \\ \vdots \\ g_m(\mathbf{x}) = g_m(x_1, x_2, \ldots, x_n) b_m \end{pmatrix}$$

  en notació vectorial aquestes restriccions s'escriuen

  $$\mathbf{x} \mathbf{0}, \quad \mathbf{g}(\mathbf{x}) \mathbf{b}$$

  on $\mathbf{0}$ és un vector columna de zeros i $\mathbf{g}(\mathbf{x})$ i $\mathbf{b}$ són els vectors columna m-dimensionals com (6). Les restriccions $g_i(\mathbf{x}),\ i = 1, 2, \ldots, m$, les suposem, com abans, contínuament diferenciables i les constants $b_i,\ i = 1, 2, \ldots, m$, són m números reals.

  Així doncs, el problema de la programació no lineal consisteix en maximitzar una certa funció escollint variables no negatives i subjecte a restriccions de desigualtat:

  $$\max_{\mathbf{x}} F(\mathbf{x}) \quad subjecte \ \ a \quad \mathbf{g}(\mathbf{x}) \mathbf{b},\ \mathbf{x} \mathbf{0}$$

- En la *programació lineal* la funció objectiu es de tipus lineal:

  $$F(\mathbf{x}) = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = \mathbf{c}\mathbf{x}$$

  on **c** és el vector fila de n constant donades:

  $$\mathbf{c} = \big(c_1, c_2, \ldots, c_n\big)$$

  i les restriccions són de dos tipus: restriccions de desigualtat lineals:

  $$\begin{pmatrix} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n b_1 \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n b_2 \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n b_m \end{pmatrix}$$

  i restriccions de no negativitat

  $$x_1 0, x_2 0, \ldots, x_n 0$$

en notació vectorial aquestes restriccions s'escriuen

$$\mathbf{x0}, \quad \mathbf{Axb}$$

on **A** és la matriu $m \times n$ exògena

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Així doncs, el problema de la programació lineal consisteix en maximitzar una certa funció lineal escollint variables no negatives i subjecte a restriccions de desigualtat lineals:

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{cx} \quad subjecte \;\; a \quad \mathbf{Axb}, \; \mathbf{x0}$$

El problema de la programació lineal és doncs un cas particular del problema de la programació no lineal en el que la funció objectiu i les restriccions són funcions lineals.

Veiem un parell d'exemples:

Exemple 2.1. *Trovar d'entre tots els rectangles de perímetre $2p > 0$, quin és el que te l'àrea més gran.*

Solució: Denotem per x i y la base i l'alada del rectangle respectivament.
El *conjunt factible* és

$$X = \{(x, y) \in \Re^2 | x \geq 0, y \geq 0, x + y = p\}$$

La *funció objectiu* és

$$f(x, y) = xy$$

Les *restriccions* són

$$\begin{aligned} x + y =& p \\ x \geq& 0 \\ y \geq& 0 \end{aligned}$$

El programa per resoldre el problema és

$$\max_{x,y} xy \quad \text{s.a} \quad x + y = p$$

Figura 2.1:

que podem transformar en

$$\max_x x(p-x)$$

i que té com solució $x = y = \frac{p}{2}$. Una extensió del problema consisteix en estudiar l'impacte de variacions de p. això s'anomena *anàlisi de sensibilitat*.

Exemple 2.2. *Trovar d'entre tots els triangles isòsceles de perímetre 2, quin és el que te l'àrea més gran.*

Solució: Denotem per y la base del triangle i per x les dues aristes d'igual longitud, de manera que $2x+y=2$. Recordem que l'àrea d'un triangle ve donada per la meitat del producte de la base per l'alada. Hem de calcular doncs l'alada d'un triangle isòsceles de perímetre 2.

El *conjunt factible* és

$$X = \{(x,y) \in \Re^2 | x \geq 0, y \geq 0, 2x+y=2\}$$

La *funció objectiu* és l'àrea del triangle,

$$A = 2(\frac{y}{2}h)\frac{1}{2} = \frac{yh}{2} = \frac{y}{2}\sqrt{x^2 - (\frac{y}{2})^2}$$

i el programa a resoldre,

$$\max_{x,y} \frac{y}{2}\sqrt{x^2 - (\frac{y}{2})^2} \text{ s.a } 2x+y=2$$

que poden reformular com

$$\max_y \frac{y}{2}\sqrt{(\frac{2-y}{2})^2 - (\frac{y}{2})^2} = \max_y \frac{y}{2}\sqrt{2(2-y)}$$

que te com solució $x = \frac{1}{3},\ y = \frac{4}{3}$.

Figura 2.2:

2.2 Tipus de màxims, el teorema de Weierstrass, i el teorema local-global.

En el problema general de programació matemàtica expressat a (3), diem que l'instrument **x** és un *màxim global* (o solució) si és factible i permet obtenir un valor de la funció objectiu que mai és menor que el valor que pren la funció per qualsevol altre instrument:

$$\mathbf{x}^* \in X \quad i \quad F(\mathbf{x}^*)F(\mathbf{x}) \quad \forall \mathbf{x} \in X$$

El màxim global $\mathbf{x}^*$ és un *màxim global estricte* si el valor de la funció en $\mathbf{x}^*$ és superior al valor de la funció en qualsevol altra punt:

$$F(\mathbf{x}^*) > F(\mathbf{x}) \quad \forall \mathbf{x} \in X,\ \mathbf{x} \neq \mathbf{x}^*.$$

Un màxim global estricte és naturalment únic, doncs si $\mathbf{x}^{**}$ i $\mathbf{x}^*$ fossin màxims globals diferents, aleshores això voldria dir que $F(\mathbf{x}^{**}) > F(\mathbf{x}^*)$ i també $F(\mathbf{x}^*) > F(\mathbf{x}^{**})$, el que és contradictori.

Nota 2.1. *Suposem que f és una funció de n variables definitionida sobre un conjunt S a $\Re^n$ i que $f(\mathbf{x}) \leq f(\mathbf{c}),\ \forall \mathbf{x} \in S$ de manera que $\mathbf{c}$ maximitza f sobre S. Aleshores, $-f(\mathbf{x}) \geq -f(\mathbf{c}),\ \forall \mathbf{x} \in S$. Així doncs, $\mathbf{c}$ maximitza f sobre S ssi $\mathbf{c}$ minimitza $-f$ sobre S. Podem utilitzar aquesta senzilla observació per convertir un problema de maximització en un problema de minimització i viceversa. Gràficament, i pel cas d'una funció f definitionida en una variable aquesta observació es representa en la figura 2.2.*

Nota 2.2. *Un resultat senzill pero de considerable interes en economia sovint s'enuncia de la següent manera:*

> *Maximitzar una funció és equivalent a maximitzar una transformació estrictament creixent d'aquesta funció.*

Per exemple, suposem que volem trovar tots el parells (x, y) que maximitzen $f(x, y)$ sobre un conjunt $S \subset \Re^2$. Aleshores, podem mirar de trovar aquells *parells* (x, y) *que maximitzen sobre* S *qualsevol de les següents funcions objectiu alternatives:*

$$af(x, y) + b,\ a > 0 \tag{2.1}$$

$$e^{f(x,y)} \tag{2.2}$$

$$\ln f(x, y) \tag{2.3}$$

En el cas (2.3) *s'ha de verificar que* $f(x, y) > 0$ *sobre* S*. Els punts màxims són exactament els mateixos. Naturalment els valors màxim són diferents. Veiem un exemple concret. El problema,*

$$\max e^{x^2+2xy^2-y^3} \textit{ s.a } (x, y) \in S$$

te la mateixa solució per (x, y) *que el problema*

$$\max x^2 + 2xy^2 - y^3 \textit{ s.a } (x, y) \in S$$

perque la funció $u \to e^u$ *és estrictament creixent.*

En general, és fàcil demostrar el resultat següent:

Teorema 2.1. *Sigui* $f(\mathbf{x}) = \mathbf{f}(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n})$ *una funció definitionida sobre un conjunt* $S \subset \Re^n$*, i sigui* F *una funció d'una variable definitionida sobre el recorregut de* f*. Definim* g *sobre* S *com:*

$$g(x_1, x_2, \ldots, x_n) = F(f(x_1, x_2, \ldots, x_n))$$

Aleshores,

(a) Si F *és creixent i* $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ *maximitza (minimitza)* f *sobre* S*, aleshores* $\mathbf{c}$ *també maximitza (minimitza)* g *sobre* S*.*

(b) Si F *és estrictament creixent, aleshores* $\mathbf{c}$ *maximitza (minimitza)* f *sobre* S *ssi* $\mathbf{c}$ *maximitza (minimitza)* g *sobre* S*.*

Demostració. Oferim la prova pel cas de la maximització. L'argument en el cas de minimització és paral.lel.

(a) Donat que $\mathbf{c}$ maximitza f sobre S, tenim que $f(\mathbf{x}) \leq \mathbf{f}(\mathbf{c})\ \forall \mathbf{x} \in \mathbf{S}$. Pero aleshores, $g(\mathbf{x}) = \mathbf{F}(\mathbf{f}(\mathbf{x})) \leq \mathbf{F}(\mathbf{f}(\mathbf{c})) = \mathbf{g}(\mathbf{c})\ \forall \mathbf{x} \in \mathbf{S}$, perque F és creixent. Per tant, $\mathbf{c}$ maximitza g sobre S.

(b) Si F és estrictament creixent i $f(\mathbf{x}) > \mathbf{f}(\mathbf{c})$, aleshores ha de ser veritat que $g(\mathbf{x}) = \mathbf{F}(\mathbf{f}(\mathbf{x})) > \mathbf{F}(\mathbf{f}(\mathbf{c})) = \mathbf{g}(\mathbf{c})$. En consequencia, $g(\mathbf{x}) \leq \mathbf{g}(\mathbf{c})\ \forall \mathbf{x} \in \mathbf{S}$ implica que $f(\mathbf{x}) \leq \mathbf{f}(\mathbf{c})\ \forall \mathbf{x} \in \mathbf{S}$.

□

Encara que hem vist uns exemples on la solució del programa existeix, això no és sempre el cas. Per il.lustrar els problemes que poden sorgir en l'existència de solució en els programas d'optimització considerem el següent exemple:

Exemple 2.3. *(El recorregut no és afitat).*
Solucionar $\max_x x^2$ *s.a* $x \in \Re$.

Aquest programa no te solució perque donat qualsevol $M \in \Re$ sempre podem trovar $x \in \Re$ tal que $x > M$. El problema és doncs que el conjut $f(x) = [0, \infty)$ no és afitat superiorment.

Exemple 2.4. *(El recorregut és afitat pero no te màxim).*
Solucionar $\max_x x^2$ *s.a* $x \in (0, 1)$.

Exemple 2.5. *(El recorregut no és afitat encara que el conjunt factibler ho sigui).*
Solucionar

$$\max_x f(x) = \begin{cases} \frac{1}{x} & \textit{si } x > 0, \\ 0 & \textit{si } x = 0. \end{cases} \quad \textit{s.a } x \in [0, 1]$$

En aquest cas, donat qualsevol M arbitrariament gran, sempre podem trobar un x suficientment petit tal que es verifiqui que $\dfrac{1}{x} > M$. Per tant, el recorregut no és superiorment afitat i el problema no té solució.

La questió pertinent és doncs, sota quines condicions podem *assegurar* que el problema d'optimització te solució?, o en altres paraules, quines són les *condicions suficients* que garanteixen l'existència de solució al problema d'optimització?. La resposta és el teorema de Weierstrass que es presenta a continuació. Notem que el teorema de Weierstrass és un teorema d'existència, és a dir ens diu sota quines condicions l'existència de solució està garantida pero no ens diu com podem trobar la solució.

Teorema 2.2. *Sigui $X \subset \Re^n$ un conjunt compacte i no buit, i sigui $F(\mathbf{x})$ una funció contínua en X aleshores $F(\mathbf{x})$ te un màxim global i un mínim global be en el interior del conjunt X, be a la seva frontera, en altres paraules, $\exists\ c, d \in X$ tal que $F(d) \leq F(x) \leq F(c)\ \forall x \in X$.*

Demostració. La prova es basa en que una funció contínua te una imatge compacte, és a dir, el conjunt de números reals $F(X) = \{z \in \Re^n | z = F(\mathbf{x})$ per algun $x \in X\}$ és compacte. Per definitionició, tot conjunt compacte de números reals conté la seva mínima fita superior. Si denotem per F^* aquesta fita superior mínima de $F(X)$, aleshores existeix un $\mathbf{x}^* \in X$ que satisfà $F(\mathbf{x}^*) = F^*$. Finalment, donat que $F(\mathbf{x})F(\mathbf{x}^*) \quad \forall \mathbf{x} \in X$, el punt $\mathbf{x}^*$ és un màxim global. L'argument per demostrar l'existència d'un mínim és similar. □

Figura 2.3: Solucions interiors i de cantonada en el cas unidimensional.

La figura 2.3 il.lustra el teorema de Weierstrass pel cas uni-dimensional.

Hem d'insistir que el teorema dona condicions suficients d'existència. això vol dir que si es verifiquen les condicions segur que existeix solució, pero a la vegada, pot existir solució a problemes que violin alguna de les condicions. Per exemple el problema de maximitzar la funció $F(X) = x^3$ subjecte a $0 < x1$ te una solució a $x = 1$ encara que el conjunt d'oportunitats no és compacte.

Definició 2.1. *El vector d'instruments* **x** *és un màxim local si és factible i permet obtenir un valor de la funció objectiu que mai és menor que el valor que pren la funció per qualsevol altre instrument prou proper:*

$$\mathbf{x}^* \in X \quad i \quad F(\mathbf{x}^*)F(\mathbf{x}) \quad \forall \mathbf{x} \in X \cap N_\epsilon(\mathbf{x}^*)$$

on $N_\epsilon(\mathbf{x}^*)$ és un entorn de $\mathbf{x}^*$ per un ϵ positiu arbitrariament petit.

Definició 2.2. *El màxim local $\mathbf{x}^*$ és un màxim local estricte si el valor de la funció en $\mathbf{x}^*$ és superior al valor de la funció en qualsevol altra punt prou proper:*

$$F(\mathbf{x}^*) > F(\mathbf{x}) \quad \forall \mathbf{x} \in X \cap N_\epsilon(\mathbf{x}^*),\ \mathbf{x} \neq \mathbf{x}^*.$$

Naturalment, poden haver altres màxims locals que proporcionin inclús valors superiors de la funció objectiu. Per exemple en la part superior de la figura 2.3 tant $\mathbf{x}^*$ com $\mathbf{x}^{**}$ són màxims locals, on $\mathbf{x}^{**}$ és màxim local estricte pero no màxim global.

Un segon teorema fonamental de la programació matemàtica és el *teorema local-global* que dona condicions suficients per tal que un màxim local sigui màxim global.

Teorema 2.3. *Si el conjunt d'oportunitats X és convexe i no buit, i $F(\mathbf{x})$ és una funció contínua i còncava en X, aleshores un màxim local és màxim global, i el conjunt de punts en el que s'obté el màxim és convexe. Si a més suposem que $F(\mathbf{x})$ és una funció estrictament còncava, aleshores la solució és única.*

Demostració. Donat que X és convexe, qualsevol punt que es trovi entre dos altres punts de X també pertany a X (i per tant és factible).

Figura 2.4: El teorema local-global.

Donat que $F(\mathbf{x})$ és estrictament còncava, qualsevol corda que uneix dos punts sobre $F(\mathbf{x})$ es trova per sota de la funció $F(\mathbf{x})$.

Sigui $\mathbf{x}^*$ un màxim local estricte. Considerem un punt factible $\mathbf{x^2}$ a la dreta de $\mathbf{x}^*$. aquest punt no pot ser màxim perque conectant ambdós punts amb una corda ens permet obtenir punts factibles entre ambdós, diguem-ne $\mathbf{x^1}$ tal que $F(\mathbf{x^1}) > F(\mathbf{x^2})$.

De forma paralela, per punts a l'esquerra de $\mathbf{x}^*$ podem fer el mateix argument.

En conseqüència, un màxim local estricte ha d'esser l'unic maxim global estricte. □

La figura 2.4 il.lustra aquesta situació.

2.3 Geometria del problema.

En el cas unidimensional (n=1), el problema de la programació matemàtica es pot il.lustrar geomètricament dibuixant la variable instrument, el conjunt d'oportunitats, i els valors de la funció objectiu com en les figures 2.3 i 2.4. En el cas de dues dimensions (n=2), el problema s'il.lustra representant els dos instruments x_1 i x_2 en els dos eixos, mostrant el conjunt d'oportunitats directament, i indicant la naturalesa de la funció objectiu a través de *corbes de nivell* i la *direcció de la preferència*.

Definició 2.3. *Una corba de nivell de la funció objectiu és el conjunt de punts en l'espai Euclidi pel que el valor de la funció objectiu és constant:*

$$\{\mathbf{x} \in E^n | F(\mathbf{x}) = \text{constant}\}$$

on diferentes constants donen lloc a diferentes corbes de nivell. El conjunt de corbes de nivell s'anomena el mapa de corbes de nivell.

Definició 2.4. *La direcció de la preferència és la direcció en la que el valor de la funció objectiu augmenta al ritme més ràpid. Aquesta direcció de la preferència*

Figura 2.5: Programació clàssica: solució de tangència.

Figura 2.6: Programació no lineal: solucions interior i de cantonada.

ve donada pel vector gradient de primeres derivades parcials de la funció objectiu:

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) = \Big(\frac{\partial F}{\partial x_1}(\mathbf{x}), \frac{\partial F}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial F}{\partial x_n}(\mathbf{x})\Big)$$

Geomètricament, el problema de la programació matemàtica consisteix en escollir un punt o un conjunt de punts en el conjunt d'oportunitats que assoleixin la corba de nivell associada al valor més alt possible de la funció objectiu.

La figura 2.5 il.lustra el problema de la programació clàssica, on les corbes de nivell de la funció objectiu C_k $(k = 1, 2, 3, \dots)$ augmenten en la direcció mostrada per la direcció de la preferéncia, i el conjunt d'oportunitats és la corba que passa per A' i A. En el cas que s'il.lustra la funció objectiu i la funció de restricció són no lineals, el conjunt d'oportunitats és convexe i el problema te una solució única en el punt de tangència T on la pendent de la corba de nivell iguala a la pendent del conjunt d'oportunitats.

La figura 2.6 il.lustra dos possibles solucions del problema de programació no lineal. La solució es pot trovar en un punt interior (I) o pot ser de cantonada (B).

Finalment, la figura 2.7 il.lustra dos possibles solucions del problema de programació lineal. La funció objectiu lineal dona lloc a corbes de nivell lineals, i les restriccions lineals de desigualtat i de no negativitat donen lloc al conjunt d'oportunitats sombrejat. Donat que la funció objectiu és lineal, $\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) = \mathbf{c}$, la direcció de en la que s'assoleix el màxim valor de la funció més ràpidament és la mateixa a tot arreu. En conseqüència no podem tenir una solució interior: la solució es trova be en un vèrtex (V) o al llarg de una part de la frontera del conjunt d'oportunitat (BF).

Figura 2.7: Programació lineal: solucions vertex i al llarg de la frontera.

Capítol 3

La Programació Clàssica.

El problema de la programació clàssica és el d'escollir els valors de certes variables per tal de maximitzar o minimitzar una funció donada, subjecte a un conjunt de restriccions d'igualtat:

$$\max_{x} F(\mathbf{x}) \quad subjecte\ a \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} \tag{3.1}$$

o escrit de forma extensiva

$$\begin{aligned} \max_{x_1, x_2, \ldots, x_n} \ & F(x_1, x_2, \ldots, x_n) \quad subjecte\ a \\ & g_1(x_1, x_2, \ldots, x_n) = b_1 \\ & g_2(x_1, x_2, \ldots, x_n) = b_2 \\ & \quad \ldots \\ & g_m(x_1, x_2, \ldots, x_n) = b_m \end{aligned}$$

Les n variables $x_1, x_2, \ldots, x_n$ són els *instruments* que els resumim en el vector columna **x**. La funció $F(\cdot)$ és la *funció objectiu*; les m funcions $g_1(\cdot), g_2(\cdot), \ldots, g_m(\cdot)$ són les *restriccions*, que resumim en el vector columna $\mathrm{g}(\cdot)$; finalment els paràmetres $b_1, b_2, \ldots, b_m$ són constants que resumim en el vector columna **b**.

Suposem que el número d'instruments n i el número de restriccions m són finits i a més, $n > m$. La diferència $n - m$ és el número de *graus de llibertat* del problema. També suposem que les $m+1$ funcions $F(\cdot), g_1(\cdot), g_2(\cdot), \ldots, g_m(\cdot)$ estàn donades, són continuament diferenciables, i no contenen cap element aleatori. Finalment, suposem que **b** conté números reals, i que **x** pot ser qualsevol vector real, subjecte solament a les m restriccions.

Si $n = m$ el problema es trivial. Per il.lustra-ho considerem el següent exemple: $\max_x ax^2$ s.a $bx = c$. Aquest problema el poden reformular com $\max_x a(\frac{c}{d})^2$ que naturalment al no dependre de x no pot prendre altre valor que $(\frac{c}{d})^2$.

Finalment, si $n < m$ hi dues possibilitats. O be hi ha $m - n$ restriccions redundants; o be les restriccions són inconsistents entre si, de manera que el conjunt factible es buit. Veiem-ho.

Exemple 3.1.

$$\begin{aligned} \max ax^2 \quad \textit{s.a} \quad bx &= c \\ dx^2 &= e \end{aligned}$$

Les restriccions ens diuen que $x = \dfrac{c}{b}$ *i també* $x = \sqrt{\dfrac{e}{d}}$*, de manera que poden passar dues coses; o be* $\dfrac{c}{b} = \sqrt{\dfrac{e}{d}}$ *que vol dir que una de les restriccions és redundant, o be* $\dfrac{c}{b} \neq \sqrt{\dfrac{e}{d}}$ *que vol dir que les restriccions són incompatibles i en conseqüència, el conjunt factible buit.*

Geomètricament, cada restricció

$$g_i(x_1, x_2, \ldots, x_n) = b_i, \quad i = 1, 2, \ldots, m$$

defineix un conjunt de punts en l'espai Euclidi E^n, i la intersecció dels m conjunts defineix el conjunt d'oportunitats,

$$X = \{\mathbf{x} \in E^n | \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$

El problema des d'el punt de vista geomètric és doncs trovar un punt (o un conjunt de punts) en el conjunt d'oportunitats que permeti assolir la més alta corba de nivell de la funció objectiu. Donat que la funció objectiu és continua i el conjunt d'oportunitats és tancat, pel teorema de Weierstrass sabem que existeix una solució si el conjunt d'oportunitats és, a més, afitat i no buit.

3.1 Maximització Lliure.

La maximització lliure (sense restriccions) d'una funció objectiu és un cas particular en el que $m = 0$.

(i) $n = 1$.

Si per efectes d'exposició considerem també $n = 1$, el problema que tenim és escollir un número real x que maximitzi $F(x)$. En aquest problema, si x^* és un màxim interior local, aleshores en un entorn de punts $x^* + \Delta x$, on Δx és una variació de x arbitrariament petita, tenim que

$$F(x^*) F(x^* + \Delta x) \tag{3.2}$$

Suposant que la funció $F(\cdot)$ és C^2, la funció de la dreta en l'expressió (3.2) es pot expandir en una serie de Taylor al voltant del punt x^* per obtenir,

$$F(x^* + \Delta x) = F(x^*) + \frac{dF}{dx}(x^*)\Delta x + \frac{1}{2}\frac{d^2F}{dx^2}(x^* + \theta\Delta x)(\Delta x)^2 \qquad (3.3)$$

on

$$0 < \theta < 1.$$

Substituint (3.3) a (3.2), obtenim la *desigualtat fonamental*:

$$\frac{dF}{dx}(x^*)\Delta x + \frac{1}{2}\frac{d^2F}{dx^2}(x^* + \theta\Delta x)(\Delta x)^2 0 \qquad (3.4)$$

una desigualtat que s'ha de verificar per qualsevol variació petita en l'instrument $\Delta x.$ Si Δx és positiu la desigualtat fonamental implica, dividint ambdós cantons per Δx i prenen el límit quan Δx s'aproxima a zero,

$$\frac{dF}{dx}(x^*)0.$$

Pero si Δx és negatiu, un raonament paral.lel ens diu

$$\frac{dF}{dx}(x^*)0.$$

Així doncs, la desigualtat fonamental requereix com una *condició necessària de primer ordre* que la primera derivada s'anuli en el punt màxim local:

$$\frac{dF}{dx}(x^*) = 0 \qquad (3.5)$$

Utilitzant la condició de primer ordre (3.5) en la desigualtat fonamental (3.4) obtenim, donat que $(\Delta x)^2$ és sempre positiu, que:

$$\frac{d^2F}{dx^2}(x^* + \theta\Delta x)0. \qquad (3.6)$$

Donat que (3.6) es verifica per tot $\Delta x,$ i donat que la segona derivada la hem suposat contínua, una *condició necessària de segon ordre* requereix que la segona derivada sigui no positiva en el punt màxim local:

$$\frac{d^2F}{dx^2}(x^*)0 \qquad (3.7)$$

Tenim doncs que les expressions (3.5) i (3.7) són, respectivament, les condicions necessàries de primer i segon ordre implicades per l'existència d'un màxim local en el punt x^*.

Les *condicions suficients* per que x^* sigui un màxim local estricte són que la primera derivada s'anuli i la segona derivada sigui estrictament negativa en aquest punt; és a dir, les condicions

$$\frac{dF}{dx}(x^*) = 0$$
$$\frac{d^2F}{dx^2}(x^*) < 0 \tag{3.8}$$

impliquen que x^* és un màxim local estricte:

$$F(x^*) > F(x^* + \Delta x)$$

Per demostrar la suficiencia podem fer servir la desigualtat fonamental. També de forma més directa, utilitzant el teorema del valor mig [1]. Sigui $a = x^*;\ b = x^* + \Delta x;$ i $\xi = x^* + \theta\Delta x$, aleshores el teorema del valor mig ens diu

$$\begin{aligned} F(x^* + \Delta x) - F(x^*) &= \frac{dF}{dx}(x^* + \theta\Delta x)(x^* + \Delta x - x^*) \\ &= \frac{dF}{dx}(x^* + \theta\Delta x)\Delta x \end{aligned}$$

de manera que

$$F(x^* + \Delta x) = F(x^*) + \frac{dF}{dx}(x^* + \theta\Delta x)(\Delta x) \tag{3.9}$$

on

$$0 < \theta < 1.$$

Donat que $F(x)$ és continuament diferenciable, si la seva primera derivada és zero i estrictament decreixent a x^* aleshores, si considerem punts $x^*+\Delta x$ a la dreta de x^*, i.e. $\Delta x > 0$, necessàriament ha de passar que

$$\frac{dF}{dx}(x^* + \theta\Delta x) < 0$$

mentres que si considerem punts $x^* + \Delta x$ a l'esquerra de x^*, i.e. $\Delta x < 0$, necessàriament ha de passar que

$$\frac{dF}{dx}(x^* + \theta\Delta x) > 0.$$

[1] Teorema del valor mig: si $y = f(x)$ és una funció diferenciable en l'interval $[a, b]$, existeix un punt $\xi \in [a, b]$ tal que es verifica la igualtat

$$f(b) - f(a) = f'(\xi)(b - a)$$

Veure una interpretació geomètrica a Aleksandrov, A.D. et al. pp.154-155.

Figura 3.1: Maximització lliure en una variable.

En qualsevol cas doncs,

$$\frac{dF}{dx}(x^* + \theta\Delta x)\Delta x < 0 \tag{3.10}$$

de manera que a partir de (3.9),

$$F(x^* + \Delta x) < F(x^*) \tag{3.11}$$

En termes gràfics, aquesta solució s'il.lustra a la figura 3.1.

En el punt x^* la pendent de la corva $F(x)$ és zero, i la pendent és decreixent, de manera que satisfà (3.8) i en conseqüència x^* és un màxim local estricte. Les mateixes condicions es satisfan en el punt x^{****} que també és un màxim local estricte. En els punts x^{**} i x^{***} la condició de primer ordre (pendent zero) es verifica, pero la condició de segon ordre no es verifica. La pendent és creixent a X^{**} i constant a x^{***}. El punt x^{**} és un mínim local estricte, i el punt x^{***} és un punt especial d'inflexió tant la primera com la segona derivades s'anulen. Es clar a partir de l'observació de x^{***} que la condició de primer ordre (3.5) i la condició de segon ordre (3.7) encara que necessàries, no són per elles mateixes suficients per caracteritzar un màxim.

(ii) $n > 1$

El cas d'una funció objectiu amb varios instruments es pot tractar de forma semblant. El problema és,

$$\max_{x_1, x_2, \ldots, x_n} F(x_1, x_2, \ldots, x_n)$$

Teorema 3.1. *Suposem que $f(\cdot)$ és diferenciable i que $\mathbf{x}^* \in \Re^n$ és un màxim o un mínim local de $f(\cdot)$. Aleshores,*

$$\left.\frac{\partial f(x)}{\partial x_n}\right|_{\mathbf{x}^*} = 0, \ \forall n.$$

Demostració. Suposant que existreix un màxim local en el punt $\mathbf{x}^*$,

$$F(\mathbf{x}^*)F(\mathbf{x}^* + h\Delta\mathbf{x}) \tag{3.12}$$

que en forma extensiva escribim

$$F(x_1^*, x_2^*, \ldots, x_n^*)F(x_1^* + h\Delta x_1, x_2^* + h\Delta x_2, \ldots, x_n^* + h\Delta x_n)$$

on h és un número real arbitrariament petit; Δx_j és una variació arbitraria al voltant de $x_j \;\; j = 1, 2, \ldots, n$; i $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)'$ és una direcció en E^n.

La funció en la part dreta de l'expressió (3.12) es pot considerar com una funció de h i, utilitzant una expansió de Taylor al voltant del punt $h = 0$ obtenim

$$F(\mathbf{x}^* + h\Delta\mathbf{x}) = F(\mathbf{x}^*) + h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\Delta\mathbf{x} + \frac{1}{2}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x}) \tag{3.13}$$

on $0 < \theta < 1$ i on $\frac{\partial F}{\partial \mathbf{x}}$ és el vector gradient i $\frac{\partial^2 F}{\partial \mathbf{x}^2}$ és el Hessià, és a dir

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) = \left(\frac{\partial F}{\partial x_1}(\mathbf{x}), \frac{\partial F}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial F}{\partial x_n}(\mathbf{x})\right)$$

$$\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\mathbf{x}) & \ldots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_2^2}(\mathbf{x}) & \ldots & \frac{\partial^2 F}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_n \partial x_2}(\mathbf{x}) & \ldots & \frac{\partial^2 F}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}$$

Combinant (3.12) i (3.13) obtenim la *desigualtat fonamental*:

$$h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\Delta\mathbf{x}) + \frac{1}{2}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x})0$$

que s'ha de verificar per totes les direccions $\Delta\mathbf{x}$ i per tots el números positius petits h. Dividint ambdós cantons per h i prenent el límit quan $h \to 0$, la desigualtat fonamental requereix com a condició necessària de primer ordre que el vector gradient s'anuli en el punt màxim local, i.e.

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$$

□

Figura 3.2: Punt crític que no és màxim ni mínim local.

Figura 3.3: Corbes de nivell de $f(x_1, x_2) = x_1^2 - x_2^2$.

Un màxim (mínim) local necessariament ha de ser un *punt estacionari* en el que totes les derivades parcials de primer ordre s'anulen. En altres paraules, un vector $\mathbf{x}^* \in \Re^n$ tal que el vector gradient s'anula s'anomena *punt crític*. El teorema ens diu doncs que tot màxim o mínim local és un punt crític. La implicació contraria no és certa.

Considerem, per exemple, la funció $f(x_1, x_2) = x_1^2 - x_2^2$ definida a $\Re^2$. En el origen es verifica $\bigtriangledown f(0,0) = (0,0)$ de manera que l'origen de coordenades és un punt crític, pero no és ni màxim ni mínim local d'aquesta funció. Aquesta funció s'il.lustra en les figura 3.2 i en la figura 3.3 següents.

Per caracteritzar màxims i mínims locals de $f(\cdot)$ de forma completa hem d'examinar les condicions de segon ordre.

La desigualtat fonamental aleshores requereix com condició necesària de segon ordre que el Hessià sigui definit negatiu o semidefinit negatiu[2] en el punt màxim local, és a dir

$$(\Delta\mathbf{x})' \frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^*)(\Delta\mathbf{x}) 0 \quad \forall\ \Delta\mathbf{x}.$$

Les condicions suficients per un màxim local estricte en el punt $\mathbf{x}^*$ són que $\mathbf{x}^*$ sigui un punt estacionari en el que el Hesssià sigui definit negatiu, és a

[2]Diem que un Hessià és semidefinit negatiu si els menors principals impars són no positius i els menors principals parells són no negatius. De la mateixa manera un Hessià és definit negatiu si els menors principals impars són negatius i els menors principals parells són positius.

dir les condicions

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$$
$$(\Delta \mathbf{x})' \frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^*)(\Delta \mathbf{x}) < 0$$

impliquen que $\mathbf{x}^*$ és un màxim local estricte:

$$F(\mathbf{x}^*) > F(\mathbf{x}^* + \Delta \mathbf{x})$$

Si el Hessià és negatiu semidefinit pero no negatiu definit, no podem dir que $\mathbf{x}^*$ és un màxim local. Considerem per exemple la funció $f(x) = x^3$ amb domini $\Re$. Aleshores, $Hf(0)$ és negatiu semidefinit perque $\frac{d^2 f(0)}{dx} = 0$, pero $\mathbf{x}^* = 0$ no és ni màxim local ni mínim local d'aquesta funció.

(iii) $n = 2$.

Per il.lustrar el cas de la maximització lliure, considerem el cas de dues dimensions, és a dir analitzem el problema

$$\max_{x_1, x_2} F(x_1, x_2)$$

Per tal de que un punt $\mathbf{x}^* = (x_1^*, x_2^*)'$ sigui un màxim local, la condició de primer ordre és que $\mathbf{x}^*$ sigui un punt estacionari:

$$\frac{\partial F}{\partial x_1}(\mathbf{x}^*) = 0, \quad \frac{\partial F}{\partial x_2}(\mathbf{x}^*) = 0$$

i la condició necessària de segon ordre és que el Hessià sigui definit negatiu o semidefinit negatiu en el punt màxim local. En termes dels menors principals del Hessià

$$\frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) 0 \tag{3.14}$$

$$\begin{vmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\mathbf{x}) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_2^2}(\mathbf{x}) \end{vmatrix} 0 \tag{3.15}$$

Per complertar l'anàlisi d'aquest cas, diem que un punt estacionari $\mathbf{x}^*$ és un mínim local si

$$\frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) \geq 0 \tag{3.16}$$

Figura 3.4: Màxim, mínim i punt de sella en dues dimensions sense restriccions.

i es verifica (3.15). Diem que és un punt de sella si (3.15) no es verifica. Aquests tres casos s'il.lustren a la figura 3.4 en la que els diagrames de la part superior mostren els tres casos directament, i els diagrames inferiors mostren corbes de nivell i direccions de preferència. Notem que un punt de sella representa un mímin des de la direcció x_1 i un màxim des de la direcció x_2.

3.2 Maximització amb Restriccions d'Igualtat.

Fins ara hem examinat el problema de la maximització lliure d'una funció objectiu. Ara abordarem el problema de la maximització d'una funció objectiu subjecte a restriccions d'igualtat. El *mètode dels multiplicadors de Lagrange* perque suposa una aproximació bàsica a tots els mètodes de resolució dels problemes d'optimització, pero també perque proporciona informació valuosa sobre la sensibilitat del valor òptim de la funció objectiu devant de canvis en les constants de restricció, sensibilitat que te importants interpretacions econòmiques.

(i) $n = 2,\ m = 1.$

Per introduir el mètode dels multiplicadors de Lagrange, considerem el problema

$$\max_{x_1,x_2} F(x_1,x_2) \quad \text{subjecte a} \quad \mathrm{g(x_1,x_2) = b.} \tag{3.17}$$

Suposem que existeix un màxim local $\mathbf{x}^* = (x_1^*, x_2^*)'$ i que en aquest punt una de les derivades parcials de la funció de restricció no és nula. Sense pèrdua de generalitat, considerem, doncs

$$\frac{\partial g}{\partial x_2}(\mathbf{x}^*) \neq 0$$

Donat aquest supòsit, la diferencial total

$$dg = \frac{\partial g}{\partial x_1}dx_1 + \frac{\partial g}{\partial x_2}dx_2 = 0$$

es pot escriure, en un entorn de $\mathbf{x}^*$ com

$$\frac{dx_2}{dx_1} = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \tag{3.18}$$

de manera que podem expressar x_2 com funció de x_1,

$$x_2 = h(x_1), \quad \text{on} \quad \frac{dh}{dx_1} = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \tag{3.19}$$

Ara podem escriure el problema original (3.17) com un problema de maximització lliure en la variable x_1

$$\max_{x_1} H(x_1) = F(x_1, h(x_1)).$$

Aplicant els resultats de la secció anterior, sabem que una condició de primer ordre per un màxim local és

$$\frac{dH}{dx_1} = \frac{\partial F}{\partial x_1} + \frac{\partial F}{\partial x_2}\frac{dh}{dx_1} = 0 \tag{3.20}$$

Utilitzant (3.19) podem rescriure (3.20) com

$$\begin{aligned}\frac{dH}{dx_1} &= \frac{\partial F}{\partial x_1} - \frac{\partial F}{\partial x_2}\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \\ &= \frac{\partial F}{\partial x_1} - \frac{\frac{\partial F}{\partial x_2}}{\frac{\partial g}{\partial x_2}}\frac{\partial g}{\partial x_1} = 0 \end{aligned} \tag{3.21}$$

A partir d'aqui podem seguir dos camins diferents pero equivalents:

- Definim una variable

  $$y = \frac{\frac{\partial F}{\partial x_2}}{\frac{\partial g}{\partial x_2}}$$

  de manera que podem rescriure (3.21) com

  $$\frac{\partial F}{\partial x_1} - y\frac{\partial g}{\partial x_1} = 0 \tag{3.22}$$

- A partir de (3.21) podem escriure,

  $$\frac{\frac{\partial F}{\partial x_1}}{\frac{\partial F}{\partial x_2}} = \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \tag{3.23}$$

Figura 3.5: Maximització restringida pel cas de dues variables i una restricció.

Aquesta solució en termes geomètrics es mostra a la figura 3.5.
Cada corba de nivell de F pren la forma $F(x_1, x_2) = K$ on K és una constant, de manera que a partir de la diferencial total

$$dF = \frac{\partial F}{\partial x_1} dx_1 + \frac{\partial F}{\partial x_2} dx_2 = 0$$

derivem que la pendent de la corba de nivell és

$$\left.\frac{dx_2}{dx_1}\right|_K = -\frac{\partial F / \partial x_1}{\partial F / \partial x_2}$$

Ara be, a partir de (3.18), la pendent de la funció de restricció és

$$\left.\frac{dx_2}{dx_1}\right|_{restr} = -\frac{\partial g / \partial x_1}{\partial g / \partial x_2}$$

Per tant, la condició de primer ordre (3.23) per caracteritzar un màxim, implica la solució de tangència en la que la pendent de la corba de nivell de la funció objectiu iguala a la pendent de la restricció,

$$\left.\frac{dx_2}{dx_1}\right|_K = \left.\frac{dx_2}{dx_1}\right|_{restr}$$

****

Una aproximació diferent al problema és la següent. Notem que les condicions necessàries (3.22) més la restricció original es poden obtenir com les condicions per caracteritzar un punt estacionari de la funció

$$L(x_1, x_2, y) = F(x_1, x_2) + y(b - g(x_1, x_2))$$

és a dir, les condicions

$$\frac{\partial L}{\partial x_j} = \frac{\partial F}{\partial x_j} - y \frac{\partial g}{\partial x_j} = 0 \quad \forall j = 1, 2. \tag{3.24}$$

$$\frac{\partial L}{\partial y} = b - g(x_1, x_2) = 0 \tag{3.25}$$

Notem que (3.24) és la generalització de (3.22). La variable y s'anomena "multiplicador de Lagrange", i la funció $L(\cdot,\cdot,\cdot)$ és la "funció lagrangiana"o "lagrangià".

Exemple 3.2. *Solucionar el següent programa:*

$$\max_{x_1,x_2} F(x_1,x_2) = x_1^\alpha x_2^\beta \quad \textit{s.a} \sum_{i=1}^{2} p_i x_i = m$$

Solucionarem el problema per dues vies

(a) Calculem la pendent de $F(x_1,x_2)$ i la igualem a la pendent de la funció que defineix la restricció.:

$$\begin{aligned} \left.\frac{dx_2}{dx_1}\right|_F &= -\frac{\frac{\partial F}{\partial x_1}}{\frac{\partial F}{\partial x_2}} \\ &= -\frac{\alpha x_1^{\alpha-1} x_2^\beta}{\beta x_1^\alpha x_2^{\beta-1}} \\ &= -\frac{\alpha x_2}{\beta x_1} \\ \left.\frac{dx_2}{dx_1}\right|_g &= -\frac{p_1}{p_2} \end{aligned}$$

Així doncs podem identificar un parell $(\overline{x}_1,\overline{x}_2)$ que verifiqui

$$\frac{\alpha \overline{x}_2}{\beta \overline{x}_1} = \frac{p_1}{p_2}$$

de manera que tenim un sistema

$$\begin{aligned} \alpha p_2 \overline{x}_2 - \beta p_1 \overline{x}_1 &= 0 \\ p_1 x_1 + p_2 x_2 &= m \end{aligned}$$

que te com solució

$$\begin{aligned} x_1 &= \frac{\alpha m}{(\alpha+\beta)p_1} \\ x_2 &= \frac{\beta m}{(\alpha+\beta)p_2} \end{aligned}$$

(b) Mètode dels multiplicadors de Lagrange: Formulem la funció lagrangiana com:

$$L(x_1,x_2) = x_1^\alpha x_2^\beta + y(m - p_1 x_1 - p_2 x_2)$$

Aplicant les condicions (3.22), obtenim

$$\begin{aligned}\frac{\partial L}{\partial x_1} =& \alpha x_1^{\alpha-1}x_2^{\beta} - yp_1 = 0\\ \frac{\partial L}{\partial x_2} =& \beta x_1^{\alpha}x_2^{\beta-1} - yp_2 = 0\\ \frac{\partial L}{\partial y} =& m - p_1x_1 - p_2x_2 = 0\end{aligned}$$

La solució d'aquest sistema és

$$\begin{aligned}x_1 =& \frac{\alpha m}{(\alpha+\beta)p_1}\\ x_2 =& \frac{\beta m}{(\alpha+\beta)p_2}\\ y =& \frac{\alpha^{\alpha}\beta^{\beta}m^{(\alpha+\beta-1)}}{p_1^{\alpha}p_2^{\beta}(\alpha+\beta)^{(\alpha+\beta-1)}}\end{aligned}$$

(ii) El cas general.

El problema general de la programació clàssica

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} \tag{3.26}$$

es pot tractar de forma paralela. Aplicant el mètode del multiplicador de Lagrange, el primer pas que hem de fer és introduir un vector fila de m noves variables

$$y = (y_1, y_2, \ldots, y_m)$$

denominades *multiplicadors de Lagrange*. El segon pas és definir el *lagrangià* com la suma de la funció objectiu i del producte intern del vector fila de multiplicadors de Lagrange i el vector columna que resulta de la diferencia entre els vectors de constants i de restriccions:

$$L(\mathbf{x},\mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x}))$$

que escrit de forma extensiva

$$\begin{aligned}L(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) =& F(x_1, x_2, \ldots, x_n) +\\ & \sum_{i=1}^{m} y_i(b_i - g_i(x_1, x_2, \ldots, x_n)).\end{aligned}$$

El pas final és trovar el punt $(\mathbf{x}^*, \mathbf{y}^*)$ pel que les derivades parcials de primer ordre del lagrangià s'anul.len:

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) &= \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}^* \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{0} \\ \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) &= \mathbf{b} - \mathbf{g}(\mathbf{x}^*) = \mathbf{0} \end{aligned} \tag{3.27}$$

El primer conjunt de n condicions ens diu que el vector gradient de la funció objectiu ha d'igualarse amb el producte del vector de multiplicadors de Lagrange i el Jacobià de les restriccions,

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{y}^* \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*) \tag{3.28}$$

o escrit en forma extensiva

$$\frac{\partial F}{\partial x_j}(x_1^*, x_2^*, \ldots, x_n^*) = \sum_{i=1}^{m} y_i^* \frac{\partial g_i}{\partial x_j}(x_1^*, x_2^*, \ldots, x_n^*); \; j = 1, 2, \ldots, n$$

Les m condicions restants són simplement les restriccions

$$\mathbf{b} = \mathbf{g}(\mathbf{x}^*) \tag{3.29}$$

Resolent el sistema de $m + n$ equacions en (3.27) ens dona la solució per les $m + n$ incògnites: els instruments $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)'$ i el multiplicadors de Lagrange $\mathbf{y}^* = (y_1^*, y_2^*, \ldots, y_m^*)$. Suposant certes condicions de segon ordre (que veurem a continuació) els instruments $\mathbf{x}^*$ són una solució local al problema de la programació clàssica. Això és fàcil de comprendre a partir de l'observació de que les restricions estàn satisfetes (com ens diu (3.29)) i els instruments estàn escollits per tal de maximitzar un lagrangià, que, en el punt $(\mathbf{x}^*, \mathbf{y}^*)$ és simplement el valor de la funció objectiu

$$L(\mathbf{x}^*, \mathbf{y}^*) = F(\mathbf{x}^*) \tag{3.30}$$

donat que les restriccions estàn satisfetes.

Formalment, poden enunciar el següent resultat:

Teorema 3.2. *Suposem que la funció objectiu i les restriccions del problema* (3.26) *són diferenciables i que $x^* \in C$ és un màxim local restringit. Suposem també que la matriu $M \times N$*

$$\begin{pmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_1(x^*)}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_M(x^*)}{\partial x_1} & \cdots & \frac{\partial g_M(x^*)}{\partial x_N} \end{pmatrix}$$

Figura 3.6: Il.lustració del teorema en el cas d'una restricció.

te rang M[3] *(Això s'anomena* la qualificació de les restriccions*: diu que les restriccions són linealment independents a* x^**.) Aleshores existeixen números* $y_m \in \Re$*, un per cada restricció, tal que*

$$\frac{\partial f(x^*)}{\partial x_n} = \sum_{m=1}^{M} y_m \frac{\partial g_m(x^*)}{\partial x_n} \quad \textit{per} \quad \textit{cada} \quad n = 1, \ldots, N,$$

Aquestes y_m *són números reals positius o negatius i es denominen* multiplicadors de Lagrange.

El teorema ens diu que evaluat en el màxim local restringit x^*, el gradient de la funció objectiu és una combinació lineal dels gradients de les restriccions. Gràficament, això vol dir que el vector gradient de la funció objectiu ha d'estar contingut en el conus format pels gradients de les funcions de restricció. Si el gradient de la funció objectiu no estigues en aquest conus voldria dir que podriem augmentar el valor de la funció objectiu sense deixar de satisfer les restriccions, el que seria contradictori amb el supòsit de que x^* és un màxim local restringit.

Per visualitzar el contingut del teorema considerem un problema de maximització amb dos instruments i una sola restricció:

$$\max_{(x_1,x_2)\in\Re^2} f(x_1, x_2) \quad \text{s.a.} \quad g(x_1, x_2) = a$$

Sigui $x^* = (x_1^*, x_2^*)$ una solució al problema. En la figura 3.6 hem dibuxat la restricció i la corva de nivell de la funció objectiu f per el valor $f(x^*)$. Notem que els gradients[4] $Df(x^*)$ i $Dg(x^*)$ són ortogonals a la corba de nivell de f i a la restricció respectivament. En aquest cas el teorema ens diu

$$Df(x^*) = yDg(x^*) \tag{3.31}$$

[3]El rang d'una matriu A és el número de files (columnes) linealment independents. En altres paraules, és l'ordre del major menor no nul, o de forma equivalent, és el mínim número de files (columnes) que permeten produr totes les files (columnes) de A com combinacions lineals.

[4]Per tal de facilitar la notació, considerem $Df(x^*) \equiv \frac{\partial f(x^*)}{\partial x}$ i $Dg(x^*) \equiv \frac{\partial g(x^*)}{\partial x}$

Figura 3.7: Un exemple en el que no es verifica la qualificació de les restriccions.

Perqué? En primer lloc hem de'observar que la condició (3.31) implica que $Df(x^*)$ ha de ser ortogonal a la restricció. Això es així perque $Dg(x^*)$ sempre és ortogonal a la restricció, de manera que $Df(x^*) = yDg(x^*)$ necessariament també ho de ser. Ara be, per qué $Df(x^*)$ ha de ser ortogonal a la restricció? Suposem que no ho fos. Aleshores podríem fer un desplaament a una distancia petita $\triangle x$ al llarg de la restricció de manera que encara es verifiques $Df(x^*) \triangle x > 0$; en altres paraules, encara seria possible augmentar el valor de la funció objectiu tot verificant la restricció, el que seria contradictori amb el supòsit de que $x^* = (x_1^*, x_2^*)$ sigui una solució al problema.

Verifiquem que la qualificació de les restriccions és indispensable per que el problema d'optimització tingui solució amb el següent exemple:

Exemple 3.3. *Solucionar*

$$\max_{(x_1,x_2)\in\Re^2} c_1x_1 - c_2x_2 \quad \textit{s.a.} \quad a_1x_1 + a_2x_2 = b$$

Tant la restricció com les corbes de nivel de la funció objectiu són funcions lineals que representen en la figura 3.7

El vector gradient de la restricció és Dg que és ortogonal a la restricció. El vector gradient de la funció objectiu és Df que és ortogonal a les corbes de nivell. En la figura 3.7 podem veure que no existeix solució afitada i els vectors Dg i Df són linealment independents en tot punt de la recta que representa la restricció (que és el conjunt factible). Per tant no hi ha cap punt que verifiqui les condicions de primer ordre.

En termes del teorema, donat que no es satisfà la qualificació de les restriccions, no ha de ser possible trovar un número y que permeti escriure el gradient de la funció objectiu com producte del gradient de la restricció i

Figura 3.8: Un exemple en el que es verifica la qualificació de les restriccions.

d'aquest multiplicador de Lagrange. Veiem-ho:

$$\begin{aligned}
\max_{x_1,x_2,y} L(x_1,x_2,y) =& c_1x_1 - c_2x_2 + y(b - a_1x_1 - a_2x_2) \\
\frac{\partial L}{\partial x_1} =& c_1 - a_1 y = 0 \qquad (3.32) \\
\frac{\partial L}{\partial x_2} =& -c_2 - a_2 y = 0 \qquad (3.33) \\
\frac{\partial L}{\partial y} =& b - a_1x_1 - a_2x_2 = 0
\end{aligned}$$

A partir de (3.32) and (3.33) obtenim

$$\begin{aligned}
y =& \frac{c_1}{a_1} > 0 \\
y =& \frac{-c_2}{a_2} < 0
\end{aligned}$$

una contradicció.

Exemple 3.4. *Si el programa fos*

$$\max_{(x_1,x_2)\in\Re^2} c_1x_1 + c_2x_2 \quad \textit{s.a.} \quad a_1x_1 + a_2x_2 = b$$

aleshores la situació estaria representada en la figura 3.8 i tots els punts del conjunt factible són solució del programa, i verifiquen les condicions de primer ordre perque Dg i Df són linealment dependents.

Si resolvem el lagrangià obtenim

$$\begin{aligned}
\max_{x_1,x_2,y} L(x_1,x_2,y) =& c_1x_1 + c_2x_2 + y(b - a_1x_1 - a_2x_2) \\
\frac{\partial L}{\partial x_1} =& c_1 - a_1 y = 0 \qquad (3.34) \\
\frac{\partial L}{\partial x_2} =& c_2 - a_2 y = 0 \qquad (3.35) \\
\frac{\partial L}{\partial y} =& b - a_1x_1 - a_2x_2 = 0 \qquad (3.36)
\end{aligned}$$

A partir de (3.34) i (3.35) obtenim

$$\frac{a_1}{c_1} = y = \frac{a_2}{c_2}$$

d'on obtenim

$$a_1 = \frac{c_1 a_2}{c_2}$$

que substitut a (3.36) dona,

$$x_2 = \frac{b}{a_2} - \frac{c_1}{c_2} x_1$$

que és l'expressió de la restricció. En altres paraules, la solució del sistema de condicions de primer ordre ens diu que hi ha una corba de nivell de la funció objectiu que es solapa amb la restricció del problema, i per tant tenim un continuu de solucions.

Les condicions necessàries de segon ordre ens diuen que la matriu de segones derivades (Hessià) del lagrangià en respecte als instruments:

$$\frac{\partial^2 L}{\partial \mathbf{x}^2}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 L}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 L}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 L}{\partial x_2^2}(\mathbf{x}) & \dots & \frac{\partial^2 L}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 L}{\partial x_n \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 L}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}$$

evaluada en el punt màxim local $(\mathbf{x}^*, \mathbf{y}^*)$ ha de ser definida negativa o semi-definida negativa quan està subjecte a les m condicions,

$$d\mathbf{g} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{g}^*) d\mathbf{x} = 0$$

Si aquest Hessià és definit negatiu subjecte a aquestes condicions, aleshores les condicions de primer ordre (3.28) i (3.29) són suficients per caracteritzar un màxim local.

3.3 La interpretació dels multiplicadors de Lagrange.

La solució del problema de la programació clàssica ens dona no solament un vector d'instruments $(\mathbf{x}^*)$ que maximitzen la funció objectiu, sino també un vector de multiplicadors de Lagrange $(\mathbf{y}^*)$ associats a aquells instruments. Quina informació conté aquest vector de multiplicadors de Lagrange?.

Proposició 3.1. *Els multiplicadors de Lagrange evaluats a la solució proporcionen una mesura de la sensibilitat del valor òptim de la funció objectiu $F^* = F(\mathbf{x}^*)$ a variacions en les constants de restricció* **b**, *és a dir,*

$$\mathbf{y}^* = \frac{\partial F^*}{\partial \mathbf{b}} \tag{3.37}$$

o en forma extensiva

$$y_i^* = \frac{\partial F^*}{\partial b_i}, \quad i = 1, 2, \ldots, m.$$

Demostració. La demostració te dues parts:

- Per demostrar (3.37) primer hem de demostrar que si les b's les considerem variables, aleshores és possible resoldre les x's i les y's com funcions de les b's. Per fer això considerem les condicions de primer ordre (3.27) que podem escriure com

$$\Psi^1(\mathbf{b}, \mathbf{y}, \mathbf{x}) \equiv \mathbf{b} - \mathbf{g}(\mathbf{x}) = \mathbf{0}$$
$$\Psi^2(\mathbf{b}, \mathbf{y}, \mathbf{x}) \equiv \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) - \mathbf{y}\frac{\partial g}{\partial \mathbf{x}}(\mathbf{x})$$

  un sistema de $m + n$ equacions i $2m + n$ variables $(\mathbf{b}, \mathbf{y}, \mathbf{x})$. Utilitzant el teorema de la funció implícita[5] podem resoldre el sistema de $m + n$ condicions de primer ordre pels instruments i els multiplicadors de Lagrange com funcions de les constrants **b**

$$\mathbf{y} = \mathbf{y}(\mathbf{b})$$
$$\mathbf{x} = \mathbf{x}(\mathbf{b})$$

- Considerem ara el Lagrangià expressat en funció de les constants **b** :

$$L(\mathbf{b}) = F(\mathbf{x}(\mathbf{b})) + \mathbf{y}(\mathbf{b})[\mathbf{b} - \mathbf{g}(\mathbf{x}(\mathbf{b}))]$$

  Diferenciant en respecte a **b** obtenim,

$$\begin{aligned}\frac{\partial L}{\partial \mathbf{b}} &= \frac{\partial F}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{b}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{b}} + (\mathbf{b} - \mathbf{g}(\mathbf{x}))'\frac{\partial \mathbf{y}'}{\partial \mathbf{b}} + \mathbf{y} \\ &= \Big(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\Big)\Big(\frac{\partial \mathbf{x}}{\partial \mathbf{b}}\Big) + (\mathbf{b} - \mathbf{g}(\mathbf{x}))'\frac{\partial \mathbf{y}'}{\partial \mathbf{b}} + \mathbf{y}\end{aligned}$$

[5] Teorema de la funció implícita: Si $F(x, y)$ és C^k en un conjunt A, (x_0, y_0) és un punt interior de A, $F(x_0, y_0) = c$ i $F_2^{'}(x_0, y_0) \neq 0$, aleshores $F(x, y) = c$ defineix y com una C^k-funció de x, $y = \varphi(x)$, en un entorn de (x_0, y_0), i $\frac{dy}{dx} = -\frac{F_1^{'}(x,y)}{F_2^{'}(x,y)}$. El teorema general de la funció implícita per sistemes d'equacions es trova a Berck i Sydsæter p.21.

Evaluat a la solució $(\mathbf{x}^*, \mathbf{y}^*)$, els primers dos termes s'anulen donades les condicions de primer ordre (3.27), de manera que la variació en el lagrangià s'iguala al vector de multiplicadors de Lagrange. Pero evaluat a la solució, el valor del lagrangià és el valor òptim de la funció objectiu (3.30). Així doncs,

$$\frac{\partial L}{\partial \mathbf{b}}(\mathbf{x}^*, \mathbf{y}^*) = \frac{\partial F^*}{\partial \mathbf{b}} = \mathbf{y}^*.$$

□

Per tant el mètode dels multiplicadors de Lagrange, a més de solucionar el problema de la programació clàssica, també ens proporciona una anàlisi de sensibilitat, mostrant en els valors dels multiplicadors de Lagrange quan sensible és el valor òptim de la funció objectiu a variacions en les constants de restricció. Per exemple, si un multiplicador de Lagrange qualsevol és igual a zero evaluat a la solució, aleshores variacions petites en la corresponent constant de restricció no afectarien el valor òptim de la funció objectiu, en altres paraules, evaluada a la solució la corresponent restricció no és operativa.

Els multiplicadors de Lagrange tenen una especial i important interpretació en problemes econòmics. En problemes d'assignació de recursos en els que la funció objectiu te les dimensions d'un valor, (e.g. beneficis, ingressos, costos) i les restriccions especifiquen un valor donat per una certa quantitat (e.g. input), aleshores el multiplicador de Lagrange medeix la sensibilitat d'un valor a variacions en una quantitat i per tant representa un preu, sovint anomenat *preu sombra* (del input).

Capítol 4

Programació No Lineal.

El problema de la *programació no lineal* és el d'escollir valors no negatius de certes variables per tal de maximitzar o minimitzar una funció objectiu donada, subjecte a un conjunt de restriccions expressades com desigualtats. Es a dir, el problema és

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{g}(\mathbf{x}) \le \mathbf{b}, \quad \mathbf{x} \ge \mathbf{0} \tag{4.1}$$

o escrit en forma extensiva

$$\begin{aligned} \max_{x_1, x_2, \ldots, x_n} & F(x_1, x_2, \ldots, x_n) \quad \text{s.a} \\ & g_1(x_1, x_2, \ldots, x_n) \le b_1 \\ & g_2(x_1, x_2, \ldots, x_n) \le b_2 \\ & \vdots \\ & g_m(x_1, x_2, \ldots, x_n) \le b_m \\ & x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0. \end{aligned}$$

Les n variables $x_1, x_2, \ldots, x_n$ són els *instruments,* que poden escriure de forma compacte en un vector columna **x**. La funció $F(\cdot)$ és la *funció objectiu,* i les m funcions $g_1(\cdot), g_2(\cdot), \ldots, g_m(\cdot)$ són *les funcions de restricció* que escribim de forma compacte en un vector columna $\mathbf{g}(\cdot)$. Les constants $b_1, b_2, \ldots, b_n$ són les *constant de restricció,* que de forma compacte resumim en un vector columna **b**. Suposem que m i n són finits; que les $m + 1$ funcions $F(\cdot), g_1(\cdot), g_2(\cdot), \ldots, g_m(\cdot)$ estàn donades, són contínuament diferenciables, i no contenen cap element aleatori; suposem també que **b** consisteix en un vector de números reals i que **x** pot ser qualsevol vector real, subjecte solament a les $m + n$ restriccions de (4.1).

Abans d'entrar en l'anàlisi del problema de la programació no lineal, val la pena fer alguns comentaris.

(i) No imposem cap restricció sobre els tamanys relatius de m i n, a diferència del supòsit sobre graus de llibertat en la programació clàssica.

(ii) La direcció de la desigualtat $(\leq)$ en les restriccions és solament una convenció. Per exemple, la desigualtat $x_1 - 2x_2 \geq 7$ es pot convertir en la desigualtat inversa multiplicant per -1, el que dona lloc a $-x_1 + 2x_2 \leq -7$.

(iii) Una restricció d'igualtat, per exemple $x_3 + 8x_4 = 12$ es pot substituir per dues restriccions de desigualtat $x_3 + 8x_4 \leq 12$ i $-x_3 - 8x_4 \leq -12$.

(iv) Les restriccions de no negativitat sobre els instruments no són restrictives. Si una variable particular, diguem-ne x_9 fos lliure (i.e. pogues ser positiva, negativa, o zero), aleshores la podriem substituir per la diferència entre dues variables no negatives: $x_9 = x_9^{'} - x_9^{''}$, on $x_9^{'} \geq 0$ i $x_9^{''} \geq 0$, i el problema es podria rescriure en termes d'aquestes variables.

En conseqüència, el problema de la programació clàssica (3.1) pot considerar-se com un cas particular de la programació no lineal en el que no hi ha restriccions de no negativitat i en el que les restriccions de desigualtat poden combinarse de manera que formin restriccions d'igualtat.

En termes geomètrics, cada una de les restriccions de no negativitat

$$x_j \geq 0, \quad j = 1, 2, \ldots, n$$

defineix un semi-espai de valors no negatius. La intersecció de tots aquest semi-espais és *l'ortant no negatiu*, un subconjut del espai Euclidi n-dimensional. Per exemple, en E^2 l'ortant no negatiu és el quadrant no negatiu, és a dir el primer quadrant més les seccions apropiades d'ambdós eixos.

Cada una de les restriccions de desigualtat

$$g_i(x_1, x_2, \ldots, x_n) \leq b_i, \quad i = 1, 2, \ldots, m$$

també defineix un conjunt de punts en l'espai Euclidi n-dimensional. La intersecció d'aquests m conjunts amb l'ortant no negatiu és *el conjunt d'oportunitats*

$$X = \{\mathbf{x} \in E^n | \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}\}$$

Geomètricament doncs, el problema de la programació no lineal és trovar un punt o un conjunt de punts en el conjunt d'oportunitats que permeti assolir la corba de nivell més alta de la funció objectiu. Donat que suposem que la funció objectiu és contínua i el conjunt d'oportunitats tancat, el teorema de Weierstrass ens diu que existeix una solució (màxim global) si el conjunt d'oportunitats és a més afitat i no buit. Tal solució pot trovar-se sobre la frontera o en el interior del conjunt d'oportunitats com s'ilustra en la figura 6.

Els supòsits de convexitat juguen un paper important en els problemes de programació no lineal. A partir del teorema local-global, un màxim local de la funció objectiu en (o sobre la frontera del) el conjunt factible és un màxim global i el conjunt de punts en el que un màxim global apareix és convexe si suposem que les funcions de restricció són convexes i la funció objectiu és còncava. Aquest cas també es coneix com *programació còncava*. Si a més suposem que la funció objectiu és estrictament còncava, aleshores la solució és única.

4.1 El cas de restriccions de no negativitat solament ($m = 0$).

Quan el problema de maximització no presenta restriccions de desigualtat, $m = 0$, el problema bàsic (4.1) es redueix a un problema de maximització d'una funció escollint valors no negatius del instruments:

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{x} \geq \mathbf{0} \tag{4.2}$$

Una manera d'abordar aquest problema és la tècnica utilitzada en la solució del problema de la programació clàssica sense restriccions, és a dir l'expansió de Taylor. Suposant que existex un màxim local per (4.2) a $\mathbf{x}^*$, aleshores per tots els punts en un entorn $\mathbf{x}^* + \Delta\mathbf{x}$

$$F(\mathbf{x}^*) \geq F(\mathbf{x}^* + h\Delta\mathbf{x}), \tag{4.3}$$

on $\Delta\mathbf{x}$ representa una direcció a E^n i h és un número arbitrariament petit i positiu. Suposant que $F(\mathbf{x})$ és dues vegades continuament diferenciable, la funció en el cantó dret de (4.3) es pot expandir com una serie de Taylor al voltant de $\mathbf{x}^*$ com

$$F(\mathbf{x}^* + h\Delta\mathbf{x}) = F(\mathbf{x}^*) + h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\Delta\mathbf{x} + \frac{1}{2}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})\Delta\mathbf{x},$$

on $0 < \theta < 1$.

Combinant les darreres dues equacions obtenim la *desigualtat fonamental*:

$$h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\Delta\mathbf{x} + \frac{1}{2}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})\Delta\mathbf{x} \leq 0, \tag{4.4}$$

que és una condició necessària per un màxim local a $\mathbf{x}^*$. Si $\mathbf{x}^*$ és una solució interior, i.e. $\mathbf{x}^* > 0$, aleshores la desigualtat fonamental s'ha de verificar per qualsevol direcció $\Delta\mathbf{x}$ de manera que obtenim la mateixa condició de primer ordre que en la programació clàssica, és a dir l'anulació de totes les derivades parcials de primer ordre. Suposem però que per un dels instruments $x_j^* = 0$. Suposant que totes les altres variacions són iguals a zero, la desigualtat fonamental (4.4) implica que, donat $x_j^* = 0$ la única direcció factible és aquella per la que $\Delta x_j \geq 0$: (dividint per h i prenent el $\lim_{h\to 0}$,

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*)\Delta x_j \leq 0.$$

La desigualtat fonamental requereix aleshores que com a condició de primer ordre:

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*) \leq 0 \quad \text{si} \quad \mathrm{x_j^*} = 0.$$

Per tant mentres que la primera derivada en respecte a x_j necessariament es cancela quan la solució és interior $(x_j^* > 0)$ en solucions de cantonada $(x_j^* = 0)$ la primera derivada necessariament és no positiva. Donat que o be la derivada s'anula (en una solució interior) o be l'instrument pren valor zero (en una solució de cantonada), el producte d'ambdós sempre es zero:

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*)x_j^* = 0. \tag{4.5}$$

Considerant ara les n dimensions del problema podem escriure

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\mathbf{x}^* = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(\mathbf{x}^*)x_j^* = 0.$$

Aquesta condició única sobre l'anulació de la suma dels productes implica, de fet, que cada terme da la suma s'anula (i.e. implica (4.5) per cada j) donada la restricció de no negativitat dels instruments i que les primeres derivades parcials són no positives. Per tant, un màxim local a $\mathbf{x}^*$ està caracteritzat per les $2n + 1$ condicions de primer ordre:

$$\begin{aligned} \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) &\leq 0 \\ \mathbf{x}^* &\geq \mathbf{0} \\ \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\mathbf{x}^* &= 0 \end{aligned} \tag{4.6}$$

Aquestes condicions impliquen els resultats que hem mencionat abans: cada derivada parcial de primer ordre es cancela si l'instrument corresponent és positiu, i és no positiu si l'instrument és zero

$$\begin{aligned} \frac{\partial F}{\partial x_j}(\mathbf{x}^*)x_j^* &= 0 \quad \text{si} \quad \mathrm{x_j^*} > 0 \\ \frac{\partial F}{\partial x_j}(\mathbf{x}^*)x_j^* &\leq 0 \quad \text{si} \quad \mathrm{x_j^*} = 0 \\ j &= 1, 2, \ldots, n. \end{aligned}$$

Les posibles solucions alternatives al problema pel cas unidimensional es mostren a la figura 4.1: una solució interior en la que la pendent de la funció és zero, una solució de cantonada en la que la pendent de la funció objectiu és negativa, o una solució de cantonada en la que la pendent és zero.

4.2 Les condicions de Kuhn-Tucker.

- El problema general de la programació no lineal

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0} \tag{4.7}$$

es pot analitzar utilitzant els resultats de la secció anterior. Les restriccions de desigualtat es poden convertir en restriccions d'igualtat afagint un vector de m "variables de folgança"(slack variables):

$$\mathbf{s} \equiv \mathbf{b} - \mathbf{g}(\mathbf{x}) = (s_1, s_2, \ldots, s_m)',$$

de manera que el problema es pot rescriure

$$\max_{\mathbf{x},\mathbf{s}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \tag{4.8}$$

on la no negativitat de les variables de folgança assegura que les restriccions de desigualtat es verifiquen. Si (4.8) no contingués les $m+n$ restriccions de no negativitat, aleshores recuperariem el problema de programació clàssica pel que el lagrangià seria

$$L'(x, y, s) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s})$$

on $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ és un vector de multiplicadors de Lagrange, com en la secció anterior.

Les condicions necessàries de primer ordre s'obtindrien en termes d'igualar a zero les derivades parcials de L' en respecte **x**, **y** i **s**. Ara be, donada la no negativitat de **x** i **s**, les condicions sobre les primeres derivades en respecte a les $m + n$ variables les substituim per les condicions obtingudes en la secció anterior. Així doncs, les condicions de primer ordre per caracteritzar

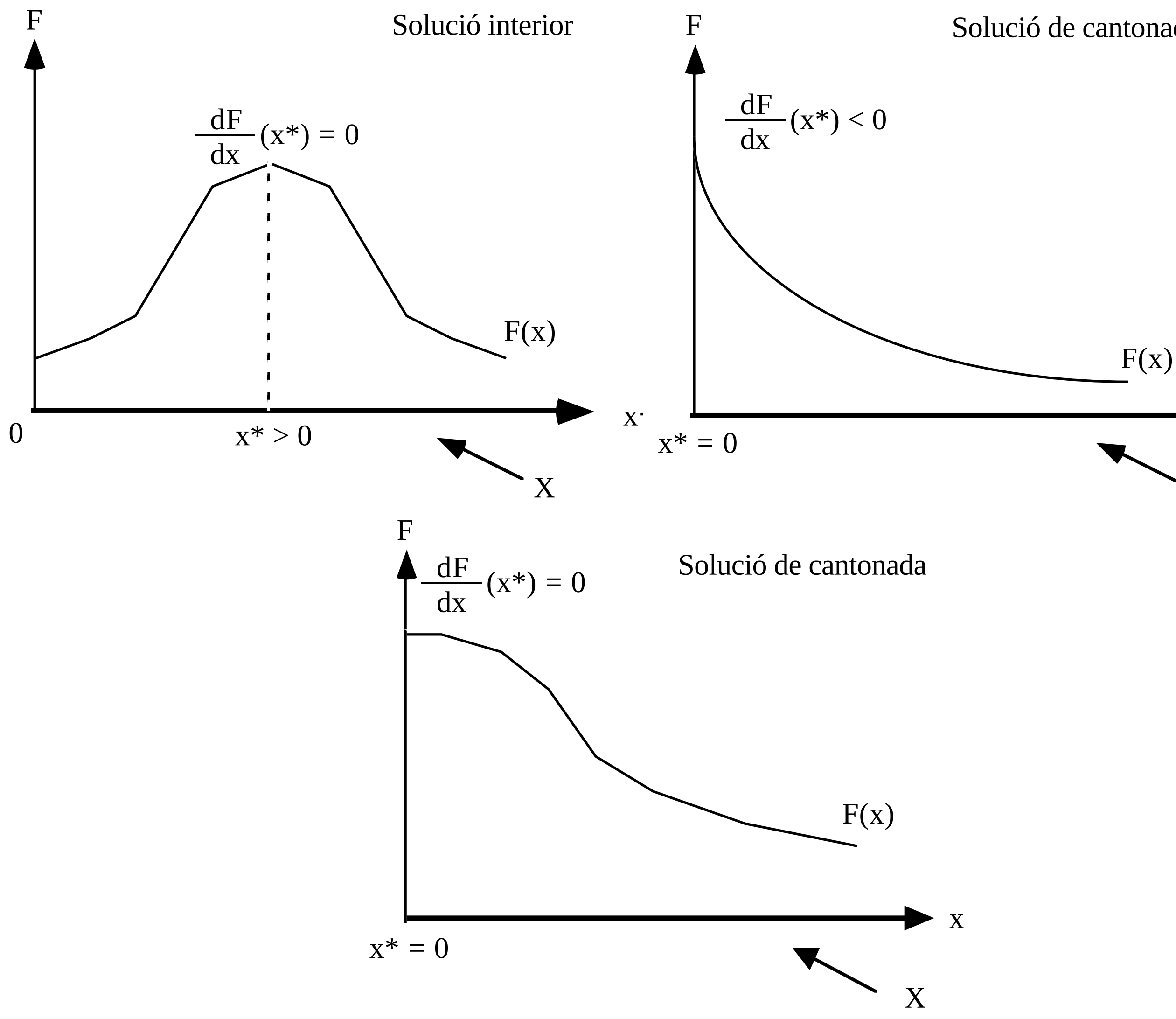

Figura 4.1: Tres possibles solucions al problema unidimensional de la maximització d'una funció objectiu restringida a valors no negatius de l'instrument.

un màxim local de (4.8) són:

$$
\begin{aligned}
\frac{\partial L'}{\partial \mathbf{x}} &= \frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} \leq \mathbf{0} \\
\frac{\partial L'}{\partial \mathbf{x}}\mathbf{x} &= \Big(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\Big)\mathbf{x} = \mathbf{0} \\
\mathbf{x} &\geq \mathbf{0} \\
\frac{\partial L'}{\partial \mathbf{y}} &= \mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s} = \mathbf{0} \qquad (4.9) \\
\frac{\partial L'}{\partial \mathbf{s}} &= -\mathbf{y} \leq \mathbf{0} \\
\frac{\partial L'}{\partial \mathbf{s}}\mathbf{s} &= -\mathbf{y}\mathbf{s} = \mathbf{0} \\
\mathbf{s} &\geq 0
\end{aligned}
$$

on totes les variables, funcions i derivades estàn evaluats a $\mathbf{x}^*,\ \mathbf{y}^*,\ \mathbf{s}^*$. Si eliminem el vector de variables de folgança $\mathbf{s}$ substituint-lo per $\mathbf{b} - \mathbf{g}(\mathbf{x})$ obtenim les *condicions de Kuhn-Tucker*:

$$
\begin{aligned}
\Big(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\Big) &\leq \mathbf{0} \\
\Big(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\Big)\mathbf{x} &= \mathbf{0} \\
\mathbf{x} &\geq \mathbf{0} \qquad (4.10) \\
\mathbf{b} - \mathbf{g}(\mathbf{x}) &\geq \mathbf{0} \\
\mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) &= \mathbf{0} \\
\mathbf{y} &\geq \mathbf{0}
\end{aligned}
$$

on totes les variables, funcions i derivades estàn evaluats a $\mathbf{x}^*,\ \mathbf{y}^*$.

- Aquestes mateixes condicions resulten si definim el lagrangià pel problema original (4.7) com[1]:

$$L = L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})),$$

[1] alternativament podem escriure el lagrangià del problema com $L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) - \mathbf{y}(\mathbf{g}(\mathbf{x}) - \mathbf{b})$

Les condicions de Kuhn-Tucker ara són

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*,\mathbf{y}^*) &= \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}^*\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*) \leq \mathbf{0}\\
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*,\mathbf{y}^*)\mathbf{x}^* &= \Big(\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}^*\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*)\Big)\mathbf{x}^* = \mathbf{0}\\
\mathbf{x}^* &\geq \mathbf{0}\\
\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*,\mathbf{y}^*) &= \mathbf{b} - \mathbf{g}(\mathbf{x}^*) \geq \mathbf{0}\\
\mathbf{y}^*\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*,\mathbf{y}^*) &= \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = \mathbf{0}\\
\mathbf{y}^* &\geq 0
\end{aligned} \tag{4.11}$$

Aquestes condicions són necessàries i suficients per un màxim local (estricte) si la funció objectiu és (estrictament) còncava i les funcions de restricció són convexes, suposant que es verifiquen unes certes condicions sobre “qualificació de les restriccions”que introduirem més endavant. Per tal de millor comprendre el significat de les condicions de Kuhn-Tucker podem expresar (4.10) de forma extensiva,

$$\frac{\partial L}{\partial x_j} = \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g}{\partial x_j} \leq 0, \qquad j = 1,2,\ldots,n \tag{4.12}$$

$$\sum_{j=1}^{n} \frac{\partial L}{\partial x_j} x_j = \sum_{j=1}^{n}\Big(\frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g}{\partial x_j}\Big) x_j = 0 \tag{4.13}$$

$$x_j \geq 0, \quad j = 1,2,\ldots,n \tag{4.14}$$

$$\frac{\partial L}{\partial y_i} = b_i - g_i(\cdot) \geq 0, \quad i = 1,2,\ldots,m \tag{4.15}$$

$$\sum_{i=1}^{m} y_i \frac{\partial L}{\partial y_i} = \sum_{i=1}^{m} y_i (b_i - g_i(\cdot)) = 0 \tag{4.16}$$

$$y_i \geq 0, \quad i = 1,2,\ldots,m \tag{4.17}$$

on totes les variables, funcions i derivades estàn evaluades a $(\mathbf{x}^*,\ \mathbf{y}^*)$.

Per entendre aquestes condicions tant importants, notem, en primer lloc que totes les restriccions de no negativitat i les restriccions de desigualtat del problema original de programació no lineal apareixen a (4.14) i (4.15) respectivament. En segon lloc notem que donat el signe de les restriccions a (4.12) i (4.14), cada terme de la suma de (4.13) ha de ser zero, de manera que

$$\text{Be } \ \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g}{\partial x_j} = 0 \ \text{ o be } \ x_j = 0 \ \text{ (o ambdós) } \ j = 1,2,\ldots,n$$

és a dir, be la condició marginal es verifica amb igualtat o l'instrument es fa zero, ó ambdós. Així:

$$\begin{aligned} \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g}{\partial x_j} \leq 0, \ \text{ pero } = 0 \ \text{ si } \ x_j^* > 0 \\ x_j^* \geq 0, \ \text{ pero } = 0 \ \text{ si } \ \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g}{\partial x_j} < 0 \\ j = 1, 2, \ldots, n \end{aligned} \tag{4.18}$$

De forma semblant, notem que donat el signe de les restriccions a (4.15) i (4.17), cada terme de la suma de (4.16) ha de ser zero, de manera que,

$$\text{Be } \ y_i = 0 \ \text{ o be } \ g_i(\mathbf{x}^*) = b_i \ \text{ (o ambdós) } \ i = 1, 2, \ldots, m$$

és a dir, be el multiplicador de Lagrange es fa zero, o la restricció de desigualtat es satisfà com igualtat estricta, ó ambdós. Així:

$$\begin{aligned} g_i(x^*) \leq b_i, \ \text{ pero } = b_i \ \text{ si } \ y_i^* > 0 \\ y_i^* \geq 0, \ \text{ pero } = 0 \ \text{ si } \ g(x^*) < b_i \\ i = 1, 2, \ldots, m \end{aligned} \tag{4.19}$$

Les condicions (4.18) i (4.19) s'anomenen *condicions de folgança complementària* (complementary slackness conditions) i són una manera alternativa de representar les condicions de Kuhn-Tucker. Finalment, com en el cas de la programació clàssica, el lagrangià evaluat a la solució és simplement el valor òptim de la funció objectiu:

$$L(\mathbf{x}^*, \mathbf{y}^*) = F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = F(\mathbf{x}^*)$$

donat que per (4.16), $\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = 0$.

LA INTERPRETACIO GEOMETRICA de les condicions de Kuhn-Tucker requereix que utilitzem la versió original de les variables de folgança (4.9) i afagim un segon vector n-dimensional no negatiu de variables de folgança,

$$\mathbf{r} = \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} - \frac{\partial F}{\partial \mathbf{x}} = (r_1, r_2, \ldots, r_n) \geq 0$$

Les condicions ara les podem expressar com

$$
\begin{aligned}
\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} + \mathbf{r} &= \mathbf{0} \\
\mathbf{r}\mathbf{x} &= \mathbf{0} \\
\mathbf{r} &\geq \mathbf{0}, \quad \mathbf{x} \geq \mathbf{0} \\
\mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s} &= \mathbf{0} \\
\mathbf{y}\mathbf{s} &= \mathbf{0} \\
\mathbf{s} &\geq \mathbf{0}, \quad \mathbf{y} \geq \mathbf{0}
\end{aligned}
$$

on totes les variables, funcions i derivades estàn evaluades a $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{r}^*, \mathbf{s}^*)$. La no negativitat de les variables de folgança assegura que les condicions de desigualtat adequades es verifiquen. El primer grup de n condicions el podem escriure com

$$
\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{y}^*\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*) + \mathbf{r}^*(-\mathbf{I})
$$

on **I** representa la matriu identitat. Geomètricament, aquestes condicions ens diuen que a la solució $\mathbf{x}^*$, el gradient de la funció objectiu, $(\frac{\partial F}{\partial \mathbf{x}})$ ha de ser una combinació ponderada dels gradients de les hipersuperficies de les restriccions, on els gradients de les restriccions de desigualtat són les files del Jacobià $(\frac{\partial \mathbf{g}}{\partial \mathbf{x}})$, els gradients de les restriccions de no negativitat són les files del negatiu de la matriu identitat, **-I**, i les ponderacions són els vectors de multiplicadors de Lagrange no negatius, $\mathbf{y}^*$, i les variables de folgança $\mathbf{r}^*$. Geomètricament, aleshores, en una solució de cantonada la direcció de preferència ha de ser una combinació lineal no negativa de les normals a la superficie que apunten cap a fora en el punt en qüestió. Aquestes normals són els vectors ortogonals $\mathbf{r}\mathbf{x} = \mathbf{0}$ i $\mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) = 0$.

Per clarificar una mica aquesta construcció, considerem l'EXEMPLE de programació no lineal següent:

$$
\begin{aligned}
\max_{x_1, x_2} F(x_1, x_2) &= -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2 \\
&\text{s.a} \\
x_1 + x_2 &\leq 1 \\
8x_1^2 + x_2^2 &\leq 2 \\
x_1 &\geq 0, \quad x_2 \geq 0
\end{aligned}
$$

on donat que la funció objectiu és estrictament còncava i les funcions de restricció són estrictament convexes, les condicions de Kuhn-Tucker i el teorema local-global, ens diuen que hi ha un únic màxim global. El lagrangià d'aquest problema és

$$\begin{aligned} L(x_1, x_2, y_1, y_2) = & -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2 + \\ & y_1(1 - x_1 - x_2) + y_2(2 - 8x_1^2 - x_2^2), \end{aligned}$$

i les condicions de Kuhn-Tucker són,

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= -16x_1 + 12x_2 - 50 - y_1 - 16y_2x_1 \leq 0 \\ \frac{\partial L}{\partial x_2} &= -20x_2 + 12x_1 + 80 - y_1 - 2y_2x_2 \leq 0 \\ \frac{\partial L}{\partial x_1}x_1 + \frac{\partial L}{\partial x_2}x_2 &= (-16x_1 + 12x_2 - 50 - y_1 - 16y_2x_1)x_1 + \\ & \quad (-20x_2 + 12x_1 + 80 - y_1 - 2y_2x_2)x_2 = 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \qquad\qquad (4.20) \\ \frac{\partial L}{\partial y_1} &= 1 - x_1 - x_2 \geq 0 \\ \frac{\partial L}{\partial y_2} &= 2 - 8x_1^2 - x_2^2 \geq 0 \\ \frac{\partial L}{\partial y_1}y_1 + \frac{\partial L}{\partial y_2}y_2 &= (1 - x_1 - x_2)y_1 + (2 - 8x_1^2 - x_2^2)y_2 = 0 \\ y_1 &\geq 0 \\ y_2 &\geq 0 \end{aligned}$$

Aquestes condicions caracteritzen una solució, pero no ens permeten trovar una solució. Per exemple, el punt

- $(x_1, x_2) = (0, 0)$ no verifica les condicions de Kuhn-Tucker, donat que en aquest punt, $(y_1, y_2) = (0, 0)$ i $\frac{\partial L}{\partial x_2} = 80 > 0$.

- $(x_1, x_2) = (\frac{1}{2}, 0)$ tampoc verifica les condicions de Kuhn-Tucker, donat que en aquest punt, $y_1 = 0$ i $\frac{\partial L}{\partial x_2} = 86 > 0$.

- $(0, 1)$ sí verifica les condicions de Kuhn-Tucker:

$$
\begin{aligned}
(x_1^*, x_2^*) &= (0, 1) \\
(y_1^*, y_2^*) &= (60, 0) \\
(\frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2})|_{\mathbf{x}^*, \mathbf{y}^*} &= (-98, 0) \\
(\frac{\partial L}{\partial y_1}, \frac{\partial L}{\partial y_2})|_{\mathbf{x}^*, \mathbf{y}^*} &= (0, 1) \\
F(x_1^*, x_2^*) &= 70 \\
(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2})|_{\mathbf{x}^*} &= (-38, 60)
\end{aligned}
$$

La representació geomètrica de la solució es mostra a la figura 4.2.

Senyalem que en la solució, la direcció de preferència (P) pren un valor intermig entre les normals (N) que apunten cap a fora.

4.3 El teorema de Kuhn-Tucker.

L'enfoc de Kuhn-Tucker al problema general de la programació no lineal:

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{subjecte a} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0} \tag{4.21}$$

com l'hem desenvolupat en la secció anterior, consisteix en introduir un vector fila de multiplicadors de Lagrange $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ on hi ha tants multipicadors com restriccions de desigualtat, i en definir el lagrangià com

$$L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})),$$

Les condicions de Kuhn-Tucker són aleshores, a partir de (4.11):

$$
\begin{aligned}
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) \leq \mathbf{0}, &\quad \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) \geq \mathbf{0} \\
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*)\mathbf{x}^* = \mathbf{0}, &\quad \mathbf{y}^* \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = \mathbf{0} \\
\mathbf{x}^* \geq \mathbf{0}, &\quad \mathbf{y}^* \geq 0
\end{aligned}
\tag{4.22}
$$

Notant la direcció de les desigualtats i recordant les condicions per un màxim, és clar que $(\mathbf{x}^*, \mathbf{y}^*)$ és un *punt de sella* del lagrangià, maximitzant-lo en respecte als instruments no negatius **x** i minimitzant-lo en respecte als multiplicadors no negatius de Lagrange **y**:

$$L(\mathbf{x}, \mathbf{y}^*) \leq L(\mathbf{x}^*, \mathbf{y}^*) \leq L(\mathbf{x}^*, \mathbf{y}) \;\; \forall\, \mathbf{x} \geq \mathbf{0} \;\; \mathbf{y} \geq \mathbf{0} \tag{4.23}$$

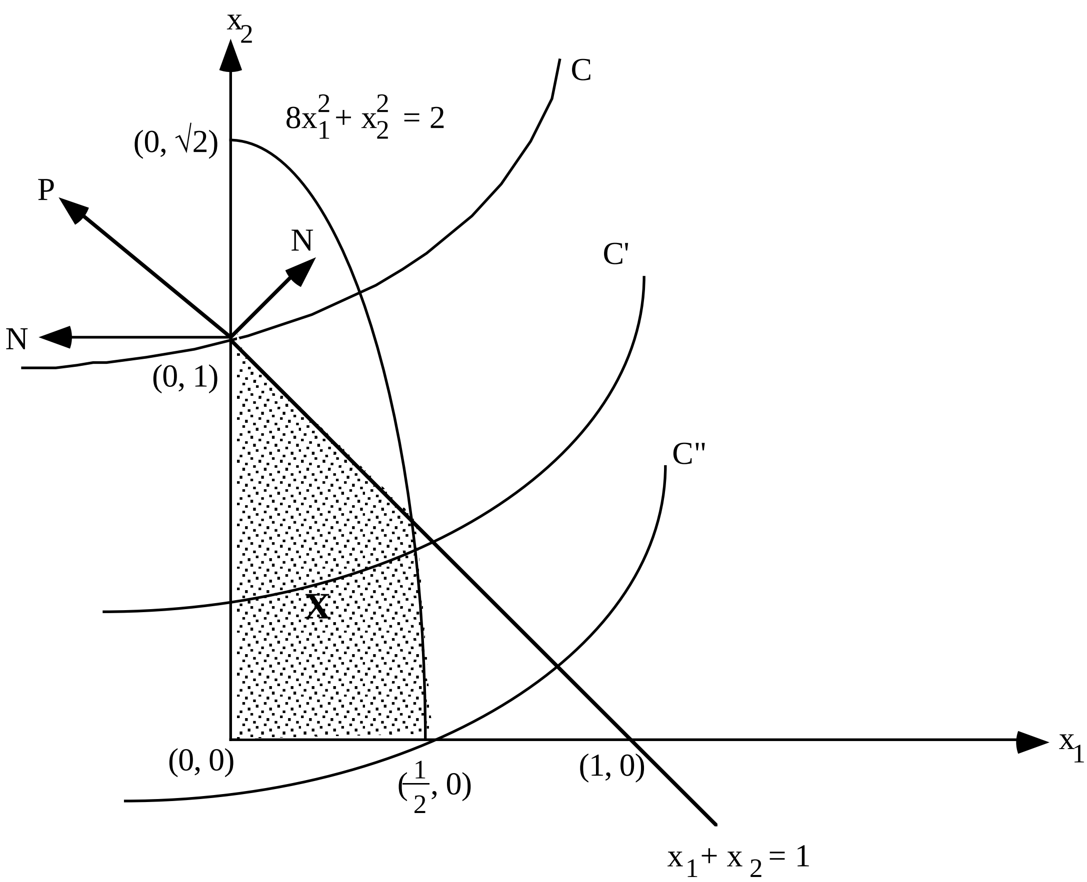

Figura 4.2: Representació geomètrica de la solució al problema de programació no lineal (4.20).

El problema de trovar vectors no negatius $(\mathbf{x}^*, \mathbf{y}^*)$ que satisfacin (4.23) es coneix com el *problema del punt de sella*.

Teorema 4.1. *El vector d'instruments* $\mathbf{x}^*$ *soluciona el problema de programació no lineal si* $(\mathbf{x}^*, \mathbf{y}^*)$ *soluciona el problema de punt de sella. També, sota certes condicions* $\mathbf{x}^*$ *soluciona el problema de programació no lineal només si existeix un vector de multiplicadors* $\mathbf{y}^*$ *tal que* $(\mathbf{x}^*, \mathbf{y}^*)$ *soluciona el problema del punt de sella.*

Demostració. (a) D'acord amb la primera part del teorema, si $(\mathbf{x}^*, \mathbf{y}^*)$ és un punt de sella com a (4.23) aleshores $\mathbf{x}^*$ soluciona el problema de la programació no lineal. Suposant que $(\mathbf{x}^*, \mathbf{y}^*)$ és tal punt de sella, donat que $\mathbf{x}^*$ maximitza el lagrangià (en respecte als instruments $\mathbf{x} \geq \mathbf{0}$):

$$F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leq F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \tag{4.24}$$

i donat que $\mathbf{y}^*$ el minimitza

$$F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \leq F(\mathbf{x}^*) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x}^*))$$

Aquesta darrera desigualtat es pot escriure com

$$(\mathbf{y} - \mathbf{y}^*)(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \geq \mathbf{0}, \quad \mathbf{y} \geq \mathbf{0} \tag{4.25}$$

i donat que els components de **y** poden ser arbitrariament grans, resulta que $\mathbf{x}^*$ ha de satisfer les restriccions de desigualtat:

$$\mathbf{g}(\mathbf{x}^*) \leq \mathbf{b}.$$

Per altra banda, escollint $\mathbf{y} = \mathbf{0}$ en (4.25), i tenint en compte que $\mathbf{y}^* \geq \mathbf{0}$ i $\mathbf{b} - \mathbf{g}(\mathbf{x}^*) \geq \mathbf{0}$ es segueix que

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = \mathbf{0}. \tag{4.26}$$

Considerem ara (4.24), que utilitzant (4.26), es pot escriure com

$$F(\mathbf{x}^*) \geq F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})), \quad \mathbf{x} \geq \mathbf{0}$$

Donat que $\mathbf{y}^*$ és no negatiu, si **x** és factible ha de ser el cas que

$$F(\mathbf{x}^*) \geq F(\mathbf{x})$$

de manera que $\mathbf{x}^*$ maximitza $F(\cdot)$ dins de la clase de **x** factibles, i per tant soluciona el problema de la programació no lineal. Aquesta demostració de la suficiencia ("si") del teorema de Kuhn-Tucker no requereix cap supòsit específic sobre les funcions $F(\cdot)$ i $g(\cdot)$.

(b) Per demostrar la part necessitat ("només si") del teorema de Kuhn-Tucker necessitem introduir alguns supòsits sobre les funcions $F(\cdot)$ i $g(\cdot)$. En particular hem de suposar que $F(\cdot)$ és una funció còncava, les funcions $\mathbf{g}(\cdot)$ són convexes, i les restriccions satisfan la *condició sobre qualificació de les restriccions* que ens diu que hi ha algun punt en el conjunt factible que satisfà totes les restriccions de desigualtat com desigualtat estricta, i.e., existeix un vector $\mathbf{x^0}$ tal que $\mathbf{x^0} \geq \mathbf{0}$ i $\mathbf{g}(\mathbf{x^0}) < \mathbf{b}$. Sota aquestes hipòtesis, suposem que $\mathbf{x}^*$ soluciona el problema de programació no lineal

$$\mathbf{x}^* \geq \mathbf{0}, \quad \mathbf{g}(\mathbf{x}^*) \leq \mathbf{b}, \text{ i } F(\mathbf{x}^*) \geq F(\mathbf{x}) \quad \forall \mathbf{x} \geq \mathbf{0}, \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}.$$

Definim ara dos conjunts en l'espai $m+1$ dimensional:

$$A = \left\{ \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \middle| \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \leq \begin{pmatrix} F(\mathbf{x}) \\ \mathbf{b} - \mathbf{g}(\mathbf{x}) \end{pmatrix} \right\} \text{ per algun } \mathbf{x} \geq \mathbf{0}$$

$$B = \left\{ \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} \middle| \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} > \begin{pmatrix} F(\mathbf{x}^*) \\ \mathbf{0} \end{pmatrix} \right\}$$

on a_0 i b_0 són escalars i **a** i **b** són vectors fila m-dimensionals. Una il.lustració d'aquests conjunts per $m = n = 1$ es mostra en la figura 4.3 on el conjunt factible és la part sombrejada de l'eix d'abcises i la solució es trova a x^*.

El conjunt A està afitat per punts amb distància vertical $F(x)$ i distància horitzontal $b - g(x)$. El conjunt B és l'interior del quadrant amb vèrtex al punt amb distància vertical $F(x^*)$ i distància horitzontal positiva. En aquest cas, i també en el cas més general, donat que $F(\cdot)$ és còncava i les funcions $g(\cdot)$ són convexes, el conjunt A és convexe. El conjunt B també és convexe donat que és l'interior d'un ortant. Donat que $\mathbf{x}^*$ soluciona el problema de programació no lineal, els dos conjunts són disjunts, de manera que pel teorema sobre hiperplans separadors per conjunts convexos disjunts, existeix un vector fila diferent de zero $(y_0, \mathbf{y})$, on y_0 és un escalar i $\mathbf{y}$ és un vector $1 \times m$ tal que:

$$(y_0, \mathbf{y}) \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \leq (y_0, \mathbf{y}) \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} \quad \forall \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \in A, \quad \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} \in B \qquad (4.27)$$

A partir de la definició de B, es segueix que $(y_0, \mathbf{y})$ és un vector no negatiu i donat que $(F(\mathbf{x}^*), \mathbf{0})'$ es trova a la frontera de B

$$y_0 F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leq y_0 F(\mathbf{x}^*) \quad \forall \mathbf{x} \geq \mathbf{0} \qquad (4.28)$$

Com a conseqüència de la condició sobre qualificació de les restriccions, $y_0 > 0$ donat que si $y_0 = 0$ aleshores la implicació de (4.28) que $\mathbf{y}(\mathbf{b} -$

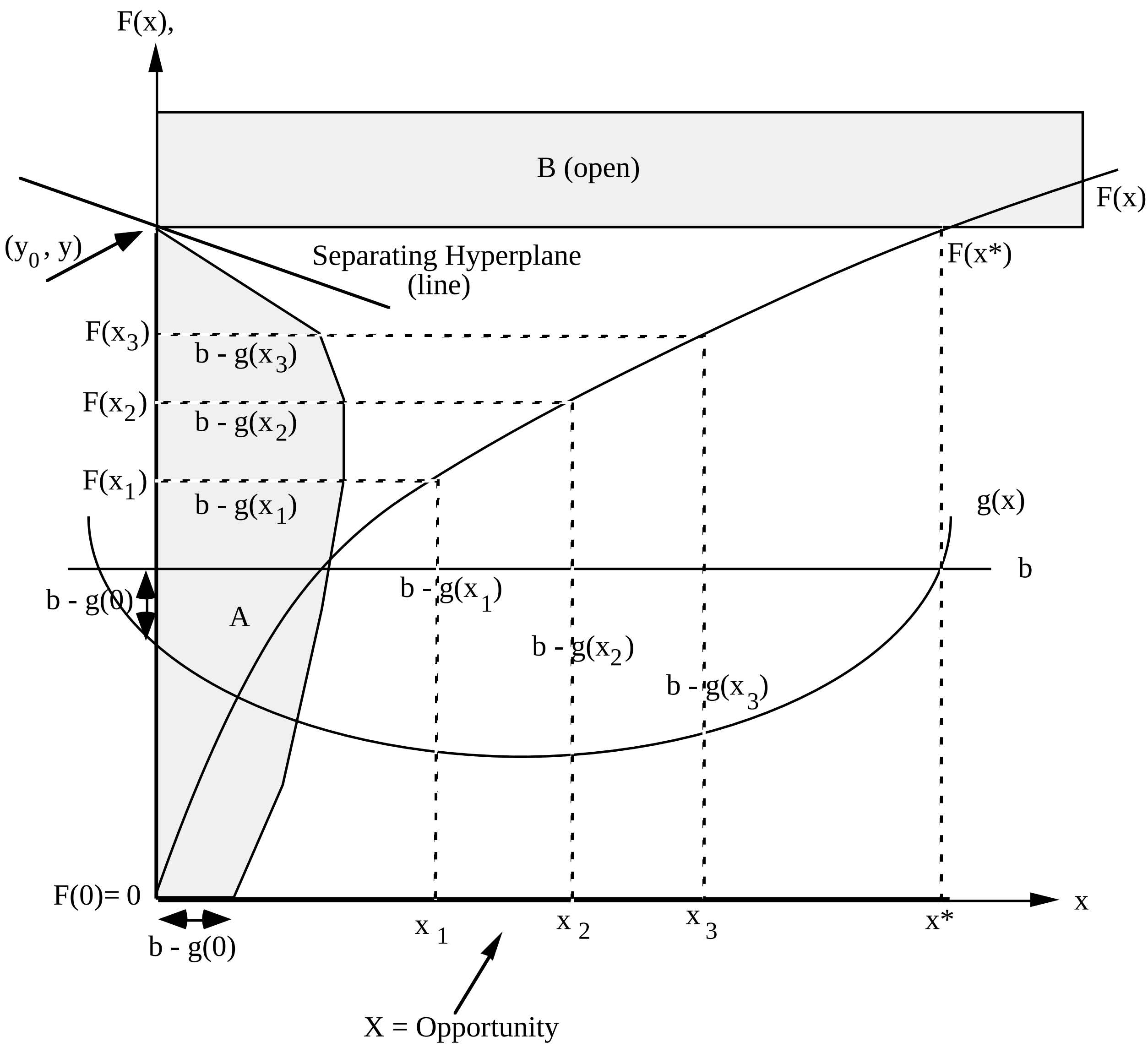

Figura 4.3: Els conjunts A i B per un problema de programació no lineal amb $m = n = 1$.

$\mathbf{g}(\mathbf{x})) \leq 0 \quad \forall \mathbf{x} \geq \mathbf{0}$ i la no negativitat de **y** contradiria l'existència d'un $\mathbf{x^0} \geq \mathbf{0}$ tal que $\mathbf{g}(\mathbf{x^0}) < \mathbf{b}$. Pero si $y_0 > 0$ aleshores ambdós cantons de (4.28) es poden dividir per y_0 per obtenir

$$F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leq F(\mathbf{x}^*) \quad \forall \mathbf{x} \geq \mathbf{0}$$
$$\text{on} \quad \mathbf{y}^* = (\frac{1}{y_0})\mathbf{y} \geq \mathbf{0}. \tag{4.29}$$

En particular, si $\mathbf{x} = \mathbf{x}^*$ aleshores

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \leq 0,$$

pero, donat que $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{b}$ i $\mathbf{y}^* \geq \mathbf{0}$

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = 0. \tag{4.30}$$

Així, definint el lagrangià com

$$L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})),$$

es segueix que a partir de (4.29), (4.30), i la no negativitat de **y**, que $(\mathbf{x}^*, \mathbf{y}^*)$ és un punt de sella per $L(\mathbf{x}, \mathbf{y})$ per $\mathbf{x} \geq \mathbf{0},\ \mathbf{y} \geq \mathbf{0}$, proporcionant d'aquesta manera la part necessària ("només si") del teorema.

Així doncs, sota els supòsits introduïts, $\mathbf{x}^*$ soluciona el problema de programació no lineal (4.21) si i solament si existeix un $\mathbf{y}^*$ tal que $(\mathbf{x}^*, \mathbf{y}^*)$ soluciona el problema del punt de sella (4.23). □

Considerem ara el problema del punt de sella amb un supòsit adicional que no hem utilitzat fins ara. Aquést és que les funcions $F(\mathbf{x})$ i $\mathbf{g}(\mathbf{x})$ són funcions diferenciables. La primera part del problema del punt de sella és la maximització de $L(\mathbf{x}, \mathbf{y}^*)$ escollint instruments **x** no negatius. Els resultats en (4.6) es poden aplicar per obtenir les següents condicions:

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) \leq \mathbf{0}$$
$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*)\mathbf{x}^* = \mathbf{0}$$
$$\mathbf{x}^* \geq \mathbf{0}$$

La segona part del problema del punt de sella, és la minimització de $L(\mathbf{x}^*, \mathbf{y})$ escollint multiplicadors de Lagrange **y** no negatius. Aquest problema genera les següents condicions:

$$\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) \geq \mathbf{0}$$
$$\mathbf{y}^*\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) = \mathbf{0}$$
$$\mathbf{y}^* \geq \mathbf{0}$$

Aquests dos conjunts de condicions són precisament les condicions de Kuhn-Tucker (4.22) que hem vist abans.

La interpretació dels multiplicadors de Lagrange.

Proposició 4.1. *Els multiplicadors de Lagrange es poden interpretar, com en el capítol anterior, com variacions en el valor òptim de la funció objectiu devant de variacions en les constants de restricció:*

$$\mathbf{y}^* = \frac{\partial F^*}{\partial b} \tag{4.31}$$

Demostració. La prova és similar a la presentada en la corresponent secció del capítol sobre programació clàssica, consistent en demostrar primer que $\mathbf{x}^*$ i $\mathbf{y}^*$ es poden solucionar com funcions de les constants de restricció i aleshores diferenciant el lagrangià en respecte a aquestes constants.

Si poguessim saber quines restriccions es satisfan com igualtats i quines com desigualtats, i quins instruments són positius i quins zero tot evaluat a la solució del problema de programació no lineal, aleshores podriem escriure les condicions de Kuhn-Tucker com igualtats. Suposem, en particular, que a la solució les restriccions es renumeren de manera que les m_1 primeres es satisfan amb igualtat i les $m - m_1$ restants es satisfan com desigualtats $(0 \leq m_1 \leq m)$ i que renumerem els instruments de manera que els n_1 primers són positius i els $n - n_1$ restants sòn zero $(0 \leq n_1 \leq n)$. Aleshores, els vectors es poden particionar com:

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} \mathbf{g^1}(\mathbf{x}) \\ \mathbf{g^2}(\mathbf{x}) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{b^1} \\ \mathbf{b^2} \end{pmatrix}, \quad \mathbf{y} = (\mathbf{y^1}, \mathbf{y^2}), \quad \mathbf{x} = \begin{pmatrix} \mathbf{x^1} \\ \mathbf{x^2} \end{pmatrix},$$

on $\mathbf{g^1}(\mathbf{x}),\ \mathbf{b^1},\ \text{i}\ \mathbf{y^1}$ consisteixen en els m_1 primers elements de $\mathbf{g}(\mathbf{x}),\ \mathbf{b},\ \text{i}\ \mathbf{y}$ respectivament, i $\mathbf{x^1}$ consisteix en els n_1 elements de $\mathbf{x}$. Les condicions de Kuhn-Tucker es poden escriure ara com

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{x^1}} &= \frac{\partial F}{\partial \mathbf{x^1}}(\mathbf{x}) - \mathbf{y^1}\frac{\partial g^1}{\partial \mathbf{x^1}}(\mathbf{x}) = \mathbf{0} \\
\mathbf{x^2} &= \mathbf{0} \\
\frac{\partial L}{\partial \mathbf{y^1}} &= \mathbf{b^1} - \mathbf{g^1}(\mathbf{x}) = \mathbf{0} \\
\mathbf{y^2} &= \mathbf{0}
\end{aligned}$$

Es clar que (4.31) es verifica pels darrers $m - m_1$ multiplicadors de Lagrange, que són iguals a zero, donat que

$$y_i^* = \frac{\partial F^*}{\partial b_i} = 0, \quad i = m_1 + 1, m_1 + 2, \ldots, m$$

Aquestes $m - m_1$ restriccions es satisfan com desigualtats, de manera que petits increments en les corresponents constants de restricció no poden fer variar el valor òptim de la funció objectiu. En respecte a les primeres m_1 multiplicadors de Lagrange, notem que el problema s'ha reduït al problema clàssic de programació

$$\max_{x^1} F(\mathbf{x^1}, \mathbf{0}) \text{ s.a. } \mathbf{g^1}(\mathbf{x^1}, \mathbf{0}) = \mathbf{b^1}$$

de manera que utilitzant el mateix argument ofert en el capítol de programació clàssica, és possible solucionar per $\mathbf{x^1}$ i $\mathbf{y^1}$ com funcions de $\mathbf{b^1}$, diferenciar el lagrangià en respecte a $\mathbf{b^1}$, i obtenir

$$y_i^* = \frac{\partial F^*}{\partial b_i} \geq 0, \quad i = 1, 2, \ldots, m_1$$

amb el que completem la prova. $\square$

4.4 Les condicions de Fritz-John.

Considerem el problema genèric de programació no lineal

$$\begin{aligned} \max_{x_1, x_2, \ldots, x_n} \; & F(x_1, x_2, \ldots, x_n) \quad \text{s.a} \\ & g_1(x_1, x_2, \ldots, x_n) \leq b_1 \\ & g_2(x_1, x_2, \ldots, x_n) \leq b_2 \\ & \quad \vdots \\ & g_m(x_1, x_2, \ldots, x_n) \leq b_m \\ & x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0. \end{aligned}$$

i expressem-lo com

$$\begin{aligned} \max_{x_1, x_2, \ldots, x_n} \; & F(x_1, x_2, \ldots, x_n) \quad \text{s.a} \\ & g_1(x_1, x_2, \ldots, x_n) - b_1 \leq 0 \\ & g_2(x_1, x_2, \ldots, x_n) - b_2 \leq 0 \\ & \quad \vdots \\ & g_m(x_1, x_2, \ldots, x_n) - b_m \leq 0 \\ & -x_1 \leq 0, -x_2 \leq 0, \ldots, -x_n \leq 0. \end{aligned}$$

de manera que se reescribim

$$\begin{aligned} & h_i(x_1, x_2, \ldots, x_n) = g_i(x_1, x_2, \ldots, x_n) - b_i, \; i = 1, \ldots, m \\ & h_j(x_1, x_2, \ldots, x_n) = -x_j, \; j = 1, \ldots, n \end{aligned}$$

podem expressar el problema original com

$$
\begin{aligned}
\max_{x_1,x_2,\ldots,x_n} & F(x_1,x_2,\ldots,x_n) \quad \text{s.a} \\
& h_1(x_1,x_2,\ldots,x_n) \leq 0 \\
& h_2(x_1,x_2,\ldots,x_n) \leq 0 \\
& \quad \vdots \\
& h_m(x_1,x_2,\ldots,x_n) \leq 0 \\
& h_{m+1}(x_1,x_2,\ldots,x_n) \leq 0 \\
& \quad \vdots \\
& h_{m+n}(x_1,x_2,\ldots,x_n) \leq 0
\end{aligned}
$$

Considerem un punt $\overline{x}^*$ tal que F i $h_i,\ i = 1,\ldots m+n$ són diferenciables en $\overline{x}^*$ i definim el conjunt $I = \{i|h_i(\overline{x}^*) = 0\}$. Aleshores si $h_i,\ i \notin I$, són continuas en $\overline{x}^*$, es verifica que quan $\overline{x}^*$ és un òptim local, existeixen escalars $\lambda_0, \lambda_i,\ i \in I$ no tots nuls tals que,

$$
\begin{aligned}
\lambda_0 \bigtriangledown F(\overline{x}^*) + \sum_{i\in I} \lambda_i \bigtriangledown g_i(\overline{x}^*) &= \overline{0} \\
\lambda_0, \lambda_i \geq 0 \text{ per } i \in I & \\
g_j(\overline{x}^*) \leq 0,\ j = 1,\ldots,m+n. &
\end{aligned}
\tag{4.32}
$$

A més, si g_i per $i \notin I$ és diferenciable en $\overline{x}^*$, aleshores les condicions (4.32) es poden expressar com,

$$
\begin{aligned}
& \lambda_0 \bigtriangledown F(\overline{x}^*) + \sum_{j=1}^{m+n} \lambda_j \bigtriangledown g_j(\overline{x}^*) = \overline{0} \\
& \lambda_j g_j(\overline{x}^*) = 0,\ j = 1,\ldots,m+n \\
& \lambda_0, \lambda_i \geq 0\ j = 1,\ldots,m+n \\
& g_j(\overline{x}^*) \leq 0,\ j = 1,\ldots,m+n.
\end{aligned}
$$

amb $\lambda_j,\ j = 1,\ldots,m+n$ no tots són nuls.

Aquestes condicions de primer ordre són solament necessàries pero no suficients per identificar un punt òptim (màxim o mínim) local.

Per demostrar que efectivament són condicions necessàries pero no suficients, considerem l'exemple següent:

$$\min_{x_1,x_2} -(x_1-5)^2-(x_2-2)^2 \quad \text{s.a}$$
$$2x_1+x_2 \leq 6$$
$$x_1 \geq 0$$
$$5 \geq x_2 \geq 0.$$

D'acord amb la figura que es mostra a continuació i que representa l'exemple que ens ocupa, el punt $\overline{x}^0 = (0,2)$ és un punt factible pero no es una solució del problema minimitzador.

INSERT FIGURE P.435 Barbolla et. al.

Ara be en aquest punt $\overline{x}^0 = (0,2)$ es verifiquen totes les condicions de Fritz-John. En efecte, la unica restricció saturada és $g_2(x_1,x_2) = -x_1 \leq 0$. Donat que $\bigtriangledown F(\overline{x}^0) = (10,0)$ i $\bigtriangledown g_2(\overline{x}^0) = (-1,0)$, existeixen $\lambda_0 \geq 0, \lambda_2 \geq 0, \lambda_1 = \lambda_3 = \lambda_4 = 0$, tals que

$$\lambda_0 \bigtriangledown F(\overline{x}^0) + \sum_{j=1}^{4} \lambda_j \bigtriangledown g_j(\overline{x}^0) = \overline{0}$$
$$\lambda_j g_j(\overline{x}^0) = 0, \; j = 1,2,3,4$$
$$\lambda_j \geq 0 \; j = 0,1,\ldots,4.$$

Solament necesitem prendre valors λ_0 i λ_2 tals que $\lambda_2 = 10\lambda_0$.

Capítol 5

La Programació Lineal.

El problema de la *Programació Lineal* consisteix en escollir valors no negatius de certes variables per tal de maximitzar o minimitzar una funció objectiu lineal donada, subjecte a un conjunt de restriccions lineals de desigualtat,

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{cx} \ \text{ s.a } \ \mathbf{Ax} \le \mathbf{b}, \quad \mathbf{x} \ge \mathbf{0}. \tag{5.1}$$

o escrit de forma extensiva,

$$\begin{aligned} &\max_{x_1, x_2, \ldots, x_n} F(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ &\text{subjecte a} \\ &\qquad a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1 \\ &\qquad a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2 \\ &\qquad \vdots \\ &\qquad a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m \\ &\qquad x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0. \end{aligned} \tag{5.2}$$

Aquest problema és un cas especial del problema de programació no lineal (4.1) en el que tant la funció objectiu com les funcions de restricció són lineals.

Les n variables $x_1, x_2, \ldots, x_n$ són els *instruments*, que poden escriure de forma compacte en un vector columna **x**. Les constants del problema són els mn *coeficients constants* $a_{11}, a_{12}, \ldots, a_{1n}; a_{21}, a_{22}, \ldots, a_{2n}; \ldots; a_{m1}, a_{m2}, \ldots, a_{mn}$ que resumim en la matriu $m \times n$ **A**; les m *constants* $b_1, b_2, \ldots, b_m$, que de forma compacta representem pel vector columna **b**; i les n *constants de la funció objectiu* $c_1, c_2, \ldots, c_n$, resumides en el vector fila **c**. Suposem que m i n són finits, que **A, b** i **c** estàn formats per numeros reals donats; i que **x** pot ser qualsevol vector real subjecte solament a les $m + n$ restriccions en (5.1).

Figura 5.1: Conjunt d'oportunitat en un problema de programació lineal amb $n = 3$ i $m = 4$..

Com en el cas de la programació no lineal, cada una de les restriccions de no negativitat

$$x_j 0, \quad j = 1, 2, \ldots, n$$

defineix un subespai tancat, i la intersecció de tots aquests subespais és l'ortant no negatiu de l'espai Euclidi E^n. Cada una de les m restriccions de desigualtat

$$H^- = \sum_{j=1}^{n} a_{ij} x_j b_i, \quad i = 1, 2, \ldots, m$$

també defineix un subespai tancat a E^n, en particular, el conjunt de punts que es troven en o en el cantó apropiat (H^-) del hiperpla corresponent definit per

$$H = \left\{ \mathbf{x} \in E^n \middle| \sum_{j=1}^{n} a_{ij} x_j = b_i \right\}$$

Un exemple són tots els punts sobre o sota un pla a E^2. En general, la intersecció de subespais tancat a E^n és un *conjunt polièdric convexe*, o si està afitat és un *poliedre convexe*. El conjunt de tots els vectors d'instruments que satisfan les $m + n$ restriccions de desigualtat i les restriccions de no negativitat de (5.1) és el conjunt factible

$$X = \{ \mathbf{x} \in E^n | \mathbf{Axb}, \quad \mathbf{x0} \}$$

Aquest conjunt d'oportunitat és per tant un conjunt polièdric convexe tancat en l'ortant no negatiu de l'espai Euclidi n-dimensional. Exemples de conjunts d'oportunitat en E^2 es mostren a la figura 7. Un exemple de conjunt d'oportunitat en E^3 es trova en la figura 5.1.

Els hiperplans afitadors s'anomenen *cares afitadores* i els punts en els que n o més cares afitadores es troven s'anomenen *vèrtices*. Cada cara afitadora consisteix en tots els punts als que una de les restriccions de desigualtat o de no negativitat es satisfà amb igualtat, i cada vèrtex és un punt en el que n o més restriccions de desigualtat es satisfan amb igualtat.

En la figura 5.1 hi ha set cares afitadores i nou vèrtices. En vuit d'aquests vèrtices, tres cares afitadores s'intersecten, i en un d'ells es produeix la conjunció de quatre cares afitadores. Els vèrtices estàn conectats per catorce eixos defint cada un d'ells per la intersecció de dues cares afitadores.

Les corves de nivell de la funció objectiu són

$$\{\mathbf{x} \in E^n | \mathbf{cx} = constant\}$$

que és l'equació d'un hiperplà a E^n. Conforme anem variant el valor de la costant, obtenim el mapa de corbes de nivell com una serie de hiperplans paralels; e.g. les línies paraleles de la figura 7. La *direcció de preferència* és la direcció de més ràpid increment de la funció objectiu i està donada pel vector gradient

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{c},$$

és a dir, un vector fila en E^n ortogonal a totes les corves de nivell per les que passa.

Geomètricament, el problema de la programació lineal consisteix en trovar un punt (o un conjunt de punts) en E^n sobre la corva de nivell de la funció objectiu que estigui el més lluny possible en la direcció de preferència pero dins del conjunt polièdric convexe d'oportunitat. A partir de la geometria és clar que si existeix una solució al problema de la programació lineal, no pot ser un punt interior sino que necessariament ha de trovar-se sobre la frontera del conjunt d'oportunitat, sobre una o més cares afitadores o de forma equivalent sobre un vèrtex, dos vèrtices, ..., n vèrtices i tots els punts entre aquests vèrtices, és a dir totes les combinacions convexes d'aquests vèrtices. La solució, quan existeix, s'obté doncs en els punt(s) en els que un hiperpla (corba de nivell) és un hiperpla suport del conjunt polièric convexe d'oportuntat.

El cas de solució única (solució en un vertex) i el cas de la solució al llarg d'una cara afitadora, solució en dos vèrtices, s'il.lustren en la figura 7. En aquest darrer cas la pendent comuna de les corbes de nivell és igual a la pendent de l'hiperpla representant la més alta cara afitadora possible, una línia en E^2, de manera que la solució es trova en dos vèrtices i també en tots el punts que els conecten. En un espai tri-dimensional $(n = 3)$ si existeix una solució, aquesta es pot trovar en un vèrtex (la intersecció de tres o mes cares afitadores), en un eix (la intersecció de dues cares afitadores) o en un pla (una cara afitadora). Mentre que la solució del problema de programació lineal, si existeix, no ha de ser necesariament única, el valor de la funció objectiu sí és únic. A més donada la convexitat del conjunt d'oportunitat i la linealitat (i.e. concavitat) de la funció objectiu, el teorema local-global ens diu que una solució que sigui màxim local és també un màxim global. En conseqüència, si en el conjunt d'oportunitat, un vèrtex proporciona un valor més alt (o amb més generalitat un valor no inferior) que els altres vèrtices al seu entorn, aleshores aquell és la solució del problema. Aquesta propietat és molt important perque constitueix la base de l'algoritme del simplex que examinarem més endavant. A més si $n > m$ aleshores les solucions ha de trovar-se en un vèrtex del conjunt d'oportunitat en el que $n - m$ o més de les variables instruments són iguals a zero; i.e., hi ha al menys una solució que te com a màxim tantes variables no zero com restriccions de desigualtat.

Donat que la funció objectiu és continua i el conjunt d'oportunitat és tancat, pel teorema de Weierstrass sabem que existeix una solució si el conjunt d'oportunitat és no buit i afitat. Per tant hi ha dues circumstàncies sota les que pot no existir solució al problema de programació lineal. La primera consisteix en que les restriccions siguin inconsistents, de manera que el conjunt d'oportunitat sigui buit. Per exemple, la restriccio $x_8 - 6$ és inconsistent amb la nonegativitat de x_8; cap punt pot satisfer simultàniament ambdues restriccions. Un altre exemple és el cas de dues restriccions de desigualtat que no tinguin cap punt en comú en l'ortant no negatiu, per exemple les restriccions $x_1 + 2x_2 6$ i $-x_1 - x_2 - 8.$

La segona circumstància en la que pot no existir solució és el cas d'un cojunt d'oportunitat no afitat i una funció objectiu que pugui augmentar sense límit en aquest conjunt. Un exemple és el problema següent:

$$\max_{x_1,x_2} x_1 + x_2 \quad \text{s.a.} \quad -\mathrm{x}_1 - \mathrm{x}_2 - 8, \quad \mathrm{x}_1 0,\ \mathrm{x}_2 0.$$

Altres exemples de les dues circumstàncies en les que no existeix solució s'il.lustren en la figura 5.2, on els punts factibles es troven en les arees sombrejades.

Si el conjunt d'oportunitat és no buit i afitat, aleshores existeix una solució que s'ha de trovar sobre la frontera del conjunt d'oportunitat. Amb més generalitat, una solució existeix si el conjunt d'oportunitat és no buit i la funció objectiu és afitada.

En general, aleshores, hi ha tres possibles solucions al problema de programació lineal: una solució única (en un vèrtex), un continu de solucions (entre dos o més vèrtices), o no existeix solució (si el conjunt d'oportunitat és buit o no afitat).

5.1 Els problemes duals de programació lineal.

Un dels fets més importants relacionats amb la programació lineal és que associat a cada problema de programació lineal hi correspon un problema dual. Si el problema original, que denominem *problema primal* és el problema de programació lineal (5.1)

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{cx} \ \text{ s.a } \ \mathbf{Ax} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \tag{5.3}$$

Figura 5.2: Dues circumstàncies en las que no existeix solució al problema de programació lineal.

aleshores, el *problema dual* és el problema de programació lineal

$$\min_{\mathbf{y}} G(\mathbf{y}) = \mathbf{y}\mathbf{b} \text{ s.a } \mathbf{y}\mathbf{A} \geq \mathbf{c}, \quad \mathbf{y} \geq \mathbf{0}. \tag{5.4}$$

on **y** és el vector fila

$$\mathbf{y} = (y_1, y_2, \ldots, y_m).$$

Escrit de forma extensiva, el problema dual és

$$\begin{aligned}
&\min_{y_1, y_2, \ldots, y_m} G(y_1, y_2, \ldots, y_m) = b_1 y_1 + b_2 y_2 + \cdots + b_m y_m \\
&\text{subjecte a} \\
&\qquad a_{11} y_1 + a_{12} y_2 + \cdots + a_{1m} y_m \geq c_1 \\
&\qquad a_{21} y_1 + a_{22} y_2 + \cdots + a_{2m} y_m c_2 \\
&\qquad \vdots \\
&\qquad a_{1n} y_1 + a_{2n} y_2 + \cdots + a_{mn} y_m \geq c_n \\
&\qquad y_1 \geq 0, y_2 \geq 0, \ldots, y_m \geq 0.
\end{aligned} \tag{5.5}$$

Les similituts entre els problemes (5.2) i (5.5) resulten evidents, Ambdós problemes consisteixen en trovar un punt extrem d'una funció objectiu lineal mitjanant l'elecció de variables no negatives subjectes a restriccions lineals de desigualtat; Ambdós problemes utilitzen el mateix conjunt de paràmetres, és a dir la matriu **A**, el vector columna **b**, i el vector fila **c**; ambdós problemes tenen un total de $m + n$ restriccions de desigualtat i ambdós problemes es poden interpretar geomètricament.

La diferència entre el problema primal i dual resideix en que el problema primal consisteix en escollir n variables resumides en un vector columna **x**, mentre que el problema dual consisteix en escollir m variables resumides en un vector fila **y**. El problema primal és un problema de maximització mentre que el problema dual és un problema de minimització; el problema original utilitza desigualtats del tipus mentre que el problema dual està sotmés a restriccions en la direcció

Figura 5.3: Els problemes duals de programació lineal en forma de taula.

oposada ; finalment, les constants de les restriccions en un problema resulten ser les constants de la funció objectiu en l'altra problema.

Si apliquem les mateixes transformacions al problema dual retrovarem el problema primal; en altres paraules, el problema primal és el dual del problema dual. Cap dels dos problemes és fonamental. Podem comenar per un qualsevol i l'altra es formula com el dual del primer, és a dir cada problema és el dual de l'altra.

Els problemes duals es poden representar en una taula com la que es mostra a la figura 5.3. El problema de programació lineal que maximitza una funció objectiu s'obté llegint d'esquerra a dreta, multipicant un element de la capsa per la corresponent variable situada a sobra de la capsa i sumant horitzontalment. El problema de programació lineal que minimitza una funció s'obté llegint de dalt cap abaix, multipicant un element de la capsa per la correspponent variable situada a l'esquerra de la capsa i sumant verticalment. L'element zero en l'extrem inferior dret de la capsa pot ser, de forma més general, qualsevol constant que es resti d'ambdues funcions objectiu.

5.2 L'enfoc lagrangià; existència, dualitat i teoremes de folgana complementaria.

La naturalesa del problema dual es pot compredre utilitzant l'analisi dels multiplicadors de Lagrange perque les variables duals poden jugar el paper de multiplicadors de Lagrange del problema primal. Suposant que el problema primal és el maximitzador,

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{cx} \;\text{ s.a }\; \mathbf{Ax} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \tag{5.6}$$

d'acord amb el teorema de Kuhn-Tucker que hem vist en l'anàlisi del problema de la programació no lineal, $\mathbf{x}^*$ és una solució a (5.6) si existeix un vector fila $\mathbf{y}^*$ tal que definint el lagrangià com

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{cx} + \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = \mathbf{cx} + \mathbf{yb} - \mathbf{yAx}$$

es verifiquen les següents condicions de Kuhn-Tucker a $\mathbf{x}^*, \mathbf{y}^*$:

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{x}} &= \mathbf{c} - \mathbf{yA0} \\
\frac{\partial L}{\partial \mathbf{x}}\mathbf{x} &= (\mathbf{c} - \mathbf{yA})\mathbf{x}^* = \mathbf{0} \\
&\mathbf{x0} \\
\frac{\partial L}{\partial \mathbf{y}} &= \mathbf{b} - \mathbf{Ax0} \\
\mathbf{y}\frac{\partial L}{\partial \mathbf{y}} &= \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = \mathbf{0} \\
&\mathbf{y}0
\end{aligned} \tag{5.7}$$

on $\mathbf{y}(\mathbf{b}-\mathbf{Ax}) = \mathbf{0}$ és la condició de folgana complementaria del problema primal. Per altra banda, si el problema primal hagués sigut el minimitzador

$$\min_{\mathbf{y}} G(\mathbf{y}) = \mathbf{yb} \text{ s.a } \mathbf{yA} \geq \mathbf{c}, \quad \mathbf{y} \geq \mathbf{0}.$$

aleshores el teorema de Kuhn-Tucker ens diu que $\mathbf{y}^*$ és solució si existeix un vector columna $\mathbf{x}^*$ tal que definint el lagrangià com

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{yb} + (\mathbf{c} - \mathbf{yA})\mathbf{x} = \mathbf{yb} + \mathbf{cx} - \mathbf{yAx}$$

es verifiquen les següents condicions de Kuhn-Tucker a $\mathbf{x}^*, \mathbf{y}^*$:

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{y}} &= \mathbf{b} - \mathbf{Ax0} \\
\mathbf{y}\frac{\partial L}{\partial \mathbf{y}} &= \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = \mathbf{0} \\
&\mathbf{y}0 \\
\frac{\partial L}{\partial \mathbf{x}} &= \mathbf{c} - \mathbf{yA0} \\
\frac{\partial L}{\partial \mathbf{x}}\mathbf{x} &= (\mathbf{c} - \mathbf{yA})\mathbf{x}^* = \mathbf{0} \\
&\mathbf{x0}
\end{aligned} \tag{5.8}$$

on $(\mathbf{c} - \mathbf{yA})\mathbf{x}^* = \mathbf{0}$ és la condició de folgana complementaria del problema dual. El lagrangià i les condicions de Kuhn-Tucker són les mateixes per ambdós problemes! El teorema fonamental de la programació lineal es basa en aquestes condicions.

El primer teorema fonamental de la programació lineal és el teorema d'existència:

Teorema 5.1 (Existència). *Una condició necessaria i suficient per l'existència d'una solució al problema primal de la programació lineal és que els conjunts factibles del problema primal i el seu dual siguin no buits. Formalment,*

$$F(\mathbf{x}^*) \geq F(\mathbf{x})\ \forall \mathbf{x} \in X \Leftrightarrow \begin{cases} X = \{\mathbf{x} \in \Re^n | \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \neq \emptyset \ i \\ Y = \{\mathbf{y} \in \Re^m | \mathbf{y}\mathbf{A} \geq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\} \neq \emptyset. \end{cases}$$

Demostració.

- ($\Leftarrow$) Per demostrar que si existeixen vectors factibles per ambdós problemes aleshores hi ha solució pels dos, considerem les restriccions de desigualtat dels problemes duals

$$\mathbf{A}\mathbf{x}\mathbf{b}$$
$$\mathbf{y}\mathbf{A}\mathbf{c}$$

Premultiplicant el primer conjut de desigualtats pel vector no negatiu **y** obtenim:

$$\mathbf{y}\mathbf{A}\mathbf{x}\mathbf{y}\mathbf{b} = G(\mathbf{y})$$

mentre que postmultiplicant el segon conjut de desigualtats pel vector no negatiu **x** obtenim:

$$F(\mathbf{x}) = \mathbf{c}\mathbf{x} \leq \mathbf{y}\mathbf{A}\mathbf{x}$$

Així doncs, si **x** i **y** són factibles,

$$F(\mathbf{x})G(\mathbf{y}) \tag{5.9}$$

és a dir el valor de la funció objectiu en el problema maximitzador no pot superar el valor de la funció objectiu del problema dual minimitzador. Suposem que existeixen vectors factibles per ambdós problemes i denotem-los per $\mathbf{x}^\circ$ i $\mathbf{y}^\circ$.

Aleshores, donat que el conjunt d'oportunitat del problema primal és no buit, conté $\mathbf{x}^\circ$, i donat que la funció objectiu és afitada

$$F(\mathbf{x})G(\mathbf{y}^\circ) \quad \forall \mathbf{x} \text{ factible},$$

es segueix que existeix una solució pel problema primal. De forma semblant, el conjunt d'oportunitat el problema dual tampoc és buit perque conté, com a mínim $\mathbf{y}^\circ$ i la funció objectiu també és afitada

$$F(\mathbf{x}^\circ)G(\mathbf{y}) \quad \forall \mathbf{y} \text{ factible},$$

de manera que el problema dual també té solució.

- ($\Rightarrow$) Per demostrar que l'existència d'una solució al problema de programació lineal implica que els conjunts d'oportunitat del problema primal i dual són no buits, suposem que $\mathbf{x}^*$ soluciona el problema maximitzador. Obviament, el problema maximitzador té un vector factible, en particular $\mathbf{x}^*$. Pero per les condicions del teorema de Kuhn-Tucker, donat que la funció objectiu és còncava i les funcions de restricció són convexes (les funcions lineals són a la vegada còncaves i convexes) i donat que es verifica la condició sobre qualificació de les restriccions, si $\mathbf{x}^*$ soluciona el problema maximitzador, aleshores existeix un vector $\mathbf{y}^*$ que satisfà les condicions (5.7), en particular

$$\mathbf{y}^*\mathbf{A} \geq \mathbf{c}, \quad \mathbf{y}^*\mathbf{0} \tag{5.10}$$

de manera que $\mathbf{y}^*$ és factible, el que complerta la prova del teorema. □

Hem demostrat doncs que existeix una solució si i només si ambdós problemes tenen vectors factibles. Si un dels problemes duals té un conjunt d'oportunitat buit aleshores l'altra be té un conjunt d'oportunitat buit o be te una funció objectiu no afitada. En general hi ha tres possibilitats pels problemes duals: ambós tenen vectors factibles i per tant pel teorema d'existència ambdós tenen solució; només un dels problemes té un vector factible, en el qual cas la seva funció objectiu no és afitada, i en conseqüència el problema de programació lineal no té solució; o be cap problema té un vector factible, el que vol dir que el conjunt d'oportunitat el problema primal és buit, i el problema de programació lineal tampoc té solució.

El segon teorema fonamental de la programació lineal és el *teorema de dualitat*.

Teorema 5.2 (Dualitat). *Una condició necessaria i suficient perque un vector, $\mathbf{x}^*$ representi una solució al problema primal de programació lineal és que existeixi un vector factible, $\mathbf{y}^*$ pel problema dual pel qual els valors de les funcions objectiu d'ambdós problemes sigui el mateix. Formalment,*

$$F(\mathbf{x}^*) \geq F(\mathbf{x})\ \forall \mathbf{x} \in X \Leftrightarrow \exists \mathbf{y}^* \in Y \ \textit{tal} \quad \textit{que} \begin{cases} G(\mathbf{y}^*) \leq G(\mathbf{y})\ \forall \mathbf{y} \in Y \ \textit{i} \\ F(\mathbf{x}^*) = G(\mathbf{y}^*) \end{cases}$$

Demostració.

- ($\Rightarrow$) Per demostrar que si $\mathbf{x}^*$ és solució del problema màxim, aleshores existeix un $\mathbf{y}^*$ que és factible pel problema dual i pel que els valors de la funció objectiu són iguals, considerem les condicions de Kuhn-Tucker (5.7). El vector $\mathbf{y}^*$ és factible perque, com hem vist a (5.10):

$$\mathbf{y}^*\mathbf{A} \geq \mathbf{c}, \quad \mathbf{y}^*\mathbf{0} \tag{5.11}$$

i les condicions

$$(\mathbf{c} - \mathbf{y}^*\mathbf{A})\mathbf{x}^* = 0$$
$$\mathbf{y}^*(\mathbf{b} - \mathbf{A}\mathbf{x}^*) = 0$$

impliquen que

$$F(\mathbf{x}^*) = \mathbf{c}\mathbf{x}^* = \mathbf{y}^*\mathbf{A}\mathbf{x}^* = \mathbf{y}^*\mathbf{b} = G(\mathbf{y}^*)$$

demostrant la igualtat dels valors de les funcions objectiu. Un raonament similar utilitzant (5.8) demostra aquesta part del teorema pel cas en el que el problema primal és un problema de minimització de la funció objectiu.

- ($\Leftarrow$) Per demostrar que si els vectors factibles existeixen per ambdós problemes pels que els valors de les funcions objectiu són iguals. aleshores aquests vectors solucionen ambdós problemes, suposem que $\mathbf{x}^*, \mathbf{y}^*$ són factibles i
$$F(\mathbf{x}^*) = \mathbf{c}\mathbf{x}^* = \mathbf{y}^*\mathbf{b} = G(\mathbf{y}^*) \tag{5.12}$$
Pero, ja sabem per (5.9) que si $\mathbf{x}, \mathbf{y}$ són factibles,
$$F(\mathbf{x})G(\mathbf{y})$$
de manera que donat que $\mathbf{y}^*$ és factible
$$F(\mathbf{x})G(\mathbf{y}^*)$$
i a partir de (5.12)
$$F(\mathbf{x})F(\mathbf{x}^*) \quad \forall \mathbf{x} \in X$$
Per tant $\mathbf{x}^*$ és solució al problema màxim. De forma paralela,
$$G(\mathbf{y}^*)G(\mathbf{y}) \quad \forall \mathbf{y} \in Y$$
el que complerta la prova.

□

Així doncs, si $\mathbf{x}^*$ és factible aleshores soluciona el problema màxim si i només si existeix un $\mathbf{y}^*$ factible pel problema dual pel que (5.12) es verifica. En particular,

$$F(\mathbf{x})F(\mathbf{x}^*) = G(\mathbf{y}^*)G(\mathbf{y})$$

Mentres F sigui més petit o igual que G, la maximització de F escollint **x** i minimitzar G escollint **y** fan augmentar el valor de F i disminuir el valor de G fins que, en la solució, ambdós valors s'igualen.

El tercer teorema fonamental de la programació lineal és el *teorema de folgana complementaria*.

Teorema 5.3 (Folgana complementaria). *Una condició necessària i suficient perque els vectors factibles* $\mathbf{x}^*, \mathbf{y}^*$ *siguin solució dels problemes duals és que satisfacin les condicions de folgana complementaria,*

$$(\mathbf{c} - \mathbf{y}^*\mathbf{A})\mathbf{x}^* = 0$$
$$\mathbf{y}^*(\mathbf{b} - \mathbf{A}\mathbf{x}^*) = 0$$

Demostració. Que sigui condició necessària es deriva directament de les condicions de Kuhn-Tucker.

La suficiencia es deriva directament del teorema de dualitat donat que, suposant que $\mathbf{x}^*, \mathbf{y}^*$ són factibles, aleshores a partir de les condicions de folgana complementaria,

$$F(\mathbf{x}^*) = \mathbf{c}\mathbf{x}^* = \mathbf{y}^*\mathbf{A}\mathbf{x}^* = \mathbf{y}^*\mathbf{b} = G(\mathbf{y}^*)$$

de manera que donat que els valors de les funcions objectiu són iguals, $\mathbf{x}^*, \mathbf{y}^*$ són solucions. □

Escrites en forma extensiva, les condicions de folgana complementaria requereixen que

$$\Big(c_j - \sum_{i=1}^{m} a_{ij}y_i^*\Big)x_j^* = 0, \quad j = 1, 2, \ldots, n$$
$$y_i^*\Big(b_i - \sum_{j=1}^{n} a_{ij}x_j^*\Big) = 0, \quad i = 1, 2, \ldots, m$$

Combinant-ho amb les restriccions de factibilitat,

$$j = 1, 2, \ldots, n \quad \begin{cases} x_j^* 0 & \text{and} = 0 \quad \text{si} \quad \sum_{i=1}^{m} a_{ij}y_i^* > c_j \\ \sum_{i=1}^{m} a_{ij}y_i^* > c_j & \text{and} = c_j \quad \text{si} \quad x_j^* > 0 \end{cases}$$
$$i = 1, 2, \ldots, m \quad \begin{cases} y_i^* 0 & \text{and} = 0 \quad \text{si} \quad \sum_{j=1}^{n} a_{ij}x_j^* < b_i \\ \sum_{j=1}^{n} a_{ij}x_j^* < b_i & \text{and} = b_i \quad \text{si} \quad y_i^* > 0 \end{cases}$$

Per tant si una certa restricció, evaluada a la solució, es satisfà com desigualtat estricta, aleshores la corresponent variable dual, evaluada a la solució és zero; i si una variable, evaluada a la solució, és positiva aleshores la corresponent restricció de desigualtat del problema dual es satisfà com igualtat. Aquestes condicions són extremadament útils en la solució de problemes de programació lineal. Per exemple, la solució del problema dual ens diu quines variables del problema primal són zero (evaluades a la solució) i quines restriccions de desigualtat del problema primal es satisfan en la solució com igualtats.

Les condicions de factibilitat i de folgana complementaria es poden expressar mitjanant variables de folgana, com vem fer en el capítol anterior, el que proporciona una important interpretació geométrica de les solucions dels problemes duals. A partir del teorema de folgana complementaria, els vectors $\mathbf{x}^*, \mathbf{y}^*$ solucionen els problemes dual de màxim i mínim si i només si:

$$\mathbf{A}\mathbf{x}^*\mathbf{b}, \quad \mathbf{x}^*\mathbf{0}, \quad \mathbf{y}^*(\mathbf{b}-\mathbf{A}\mathbf{x}^*)=0$$
$$\mathbf{y}^*\mathbf{A}\mathbf{c}, \quad \mathbf{y}^*\mathbf{0}, \quad (\mathbf{c}-\mathbf{y}^*\mathbf{A})\mathbf{x}^*=0$$

Si introduim el vector columna de variables de folgana $\mathbf{s}=(s_1, s_2, \ldots, s_m)'$ pel problema del màxim, i el vector fila de variables de folgana $\mathbf{r}=(r_1, r_2, \ldots, r_n)'$ pel problema del mínim, les condicions que caracteritzen una solució es poden rescriure com

$$\begin{aligned} &\mathbf{A}\mathbf{x}^*+\mathbf{s}^*=\mathbf{b}, \quad \mathbf{x}^*\mathbf{0}, \quad \mathbf{s}^*\mathbf{0}, \quad \mathbf{y}^*\mathbf{s}^*=0 \\ &\mathbf{y}^*\mathbf{A}=\mathbf{c}+\mathbf{r}^*, \quad \mathbf{y}^*\mathbf{0}, \quad \mathbf{r}^*\mathbf{0}, \quad \mathbf{r}^*\mathbf{x}^*=0 \end{aligned} \tag{5.13}$$

on la no negativitat de les variables de folgana assegura que les restriccions de desigualtat d'ambdós problemes es satisfan, i l'anul.lació de la suma dels productes de les variables de folgana amb les varibles duals assegura que les condicions de folgana complementaria també es satisfan.

Considerem ara la interpretació geomètrica de les condicions (5.13). Les condicions de factibilitat el problema de mínim es pot escriure

$$\mathbf{c}=\mathbf{y}^*\mathbf{A}+\mathbf{r}^*(-\mathbf{I}), \quad \mathbf{y}^*\mathbf{0}, \quad \mathbf{r}^*\mathbf{0} \tag{5.14}$$

on **I** representa la matriu identitat.

D'acord amb aquestes condicions, en la solució el vector **c** és una combinació lineal no negativa de les files de la matriu de coeficients i de les files de la matriu identitat. Pero **c** és el vector gradient de la funció objectiu pel problema màxim i per tant punts en la direcció de preferència, mentres que les files de la matriu de coeficients i de la matriu identitat amb signe negatiu són els vectors gradients per les restriccions de desigualtat i de no negativitat respectivament, que gràficament representen les normals al conjunt d'oportuntat que apunten cap a fora. De forma semblant, les condicions de factibilitat del problema màxim

$$-\mathbf{b}=(-\mathbf{A})\mathbf{x}^*+(-\mathbf{I})\mathbf{s}^*, \quad \mathbf{x}^*\mathbf{0}, \quad \mathbf{s}^*\mathbf{0} \tag{5.15}$$

ens diuen que la direcció de preferència pel problema mínim [el negatiu del vector gradient de $G(\mathbf{y})$, donat que el problema és de minimització] és una combinació lineal no negativa de les normals al conjunt d'oportuntat que apunten cap a fora, com ens donen les columnes de les matrius de coeficient negatius i la matriu identitat amb el signe negatiu. Geomètricament, (5.14) i (5.15) ens diuen que en la

Figura 5.4: Geometria del problema dual per $m = n = 2$.

solució de qualsevol dels problemes la direcció de preferència s'ha de trovar entre les normals al conjunt d'oportuntat que apunten cap a fora. Aquesta interpretació geomètrica s'il.lustra en la figura 15 pels problemes duals en els que $m = m = 2$. en ambdós problemes la direcció de preferència, P, es trova entre les normals al conjunt d'oportuntat que apunten cap a fora, N, en el punt que és solució.

Les altres condicions de (5.13) que es refereixen a la folgana complementaria:

$$\mathbf{y}^*\mathbf{s}^* = \mathbf{0}, \quad \mathbf{r}^*\mathbf{x}^* = \mathbf{0}$$

geomètricament ens diuen que no es dona cap pès a la normal que apunta cap a fora corresponent a una restricció de desigualtat (no negativitat) d'un problema si es dona un pès positiu a la normal al conjunt d'oportuntat que apunta cap a fora per la corresponent restricció de desigualtat (no negativitat) del problema dual:

$$\begin{array}{ll} y_i^* = 0 \ \text{ si } \ \mathrm{s}_\mathrm{i}^* > 0; & \mathrm{x}_\mathrm{j}^* = 0 \ \text{ si } \ \mathrm{r}_\mathrm{j}^* > 0 \\ r_j^* = 0 \ \text{ si } \ \mathrm{x}_\mathrm{j}^* > 0; & \mathrm{s}_\mathrm{i}^* = 0 \ \text{ si } \ \mathrm{y}_\mathrm{i}^* > 0, \\ i = 1, 2, \ldots, m; & j = 1, 2, \ldots, n. \end{array}$$

5.3 La interpretació de les varibles duals i l'anàlisi de sensibilitat.

Donat que les variables del problema dual són multiplicadors de Lagrange del problema primal,es poden interpretar com la sensibilitat del valors òptims de la funció objectiu en respecte a variacions en les constants de restricció.

Suposem que a la solució del problema dual (5.3) i (5.4), m_1 de les m restriccions de desigualtat del problema maximitzador són satisfetes com igualtats, mentre que $(m-m_1)$ es satisfan com desigualtats estrictes. Paralelament, n_1 de les n restriccions de desigualtat del problema minimitzador es satisfan com igualtats, i $(n - n_1)$ es satisfan con desigualtats estrictes. Poden renumerar les restriccions si s'escau, per tal de que les m_1 restriccions del problema maximitzador i les n_1 restriccions del problema minimitzador són les satisfetes com igualtats. Aquesta renumeració requereix una renumeració similar de les variables en el problema dual. La matriu de coeficients i els vectors fila i columna de paràmetres es pot particionar com,

$$\mathbf{c} = (\mathbf{c^1}, \mathbf{c^2})$$
$$\mathbf{A} = \begin{pmatrix} \mathbf{A^{11}} & \mathbf{A^{12}} \\ \mathbf{A^{21}} & \mathbf{A^{22}} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{A^1} \\ \mathbf{A^2} \end{pmatrix} \tag{5.16}$$

on $\mathbf{c^1}$ conté n_1 elements, $\mathbf{b^1}$ conté m_1 elements, i $\mathbf{A^{11}}$ conté $m_1 n_1$ elements. Els vectors columna i fila de variables es poden particionar de forma semblant,

$$\mathbf{y} = (\mathbf{y^1}, \mathbf{y^2}), \quad \mathbf{x} = \begin{pmatrix} \mathbf{x^1} \\ \mathbf{x^2} \end{pmatrix} \tag{5.17}$$

on $\mathbf{x^1}$ conté n_1 elements i $\mathbf{y^1}$ conté m_1 elements. Amb els supòsits anteriors, a la solució dels problemes duals $\mathbf{x}^*, \mathbf{y}^*$

$$\begin{aligned} \mathbf{A^{11}}\mathbf{x^{1^*}} + \mathbf{A^{12}}\mathbf{x^{2^*}} &= \mathbf{b^1} \\ \mathbf{A^{21}}\mathbf{x^{1^*}} + \mathbf{A^{22}}\mathbf{x^{2^*}} &= \mathbf{b^2} \\ \mathbf{y^{1^*}}\mathbf{A^{11}} + \mathbf{y^{2^*}}\mathbf{A^{21}} &= \mathbf{c^1} \\ \mathbf{y^{1^*}}\mathbf{A^{12}} + \mathbf{y^{2^*}}\mathbf{A^{22}} &> \mathbf{c^2} \end{aligned} \tag{5.18}$$

Aleshores, a partir dels resultats sobre folgana complementaria obtinguts en la secció anterior,

$$\begin{aligned} \mathbf{x^{1^*}} 0, \quad \mathbf{x^{2^*}} &= 0 \\ \mathbf{y^{1^*}} 0, \quad \mathbf{y^{2^*}} &= 0, \end{aligned}$$

de manera que les igualtats de (5.18) es poden escriure

$$\begin{aligned} \mathbf{A^{11}}\mathbf{x^{1^*}} &= \mathbf{b^1} \\ \mathbf{y^{1^*}}\mathbf{A^{11}} &= \mathbf{c^1} \end{aligned} \tag{5.19}$$

Les solucions per ambdós problemes existeixen i són úniques si $\mathbf{A^{11}}$ és una matriu cuadrada i no singular (en la solució vèrtex), i en tal cas,

$$\begin{aligned} \mathbf{x^{1^*}} &= (\mathbf{A^{11}})^{-1}\mathbf{b^1}, \quad \mathbf{x^{2^*}} = \mathbf{0} \\ \mathbf{y^{1^*}} &= \mathbf{c^1}(\mathbf{A^{11}})^{-1}, \quad \mathbf{y^{2^*}} = \mathbf{0} \end{aligned} \tag{5.20}$$

el que mostra les solucions a ambdós problemes explícitament en termes dels $m_1 n_1 + m_1 + n_1$ paràmetres en $\mathbf{A^{11}},\ \mathbf{b^1}$ i $\mathbf{c^1}$. Els corresponents valors òptims de les funcions objectiu són

$$F(\mathbf{x}^*) = \mathbf{c^1}\mathbf{x^{1^*}} = \mathbf{c^1}(\mathbf{A^{11}})^{-1}\mathbf{b^1} = \mathbf{y^{1^*}}\mathbf{b^1} = G(\mathbf{y}^*) \tag{5.21}$$

que mostra els valors òptims d'ambdós problemes, iguals entre si pel teorema de dualitat, explícitament en funció dels $m_1n_1 + m_1 + n_1$ paràmetres del problema.

L'anàlisi de sensibilitat tracta dels efectes que variacions en els paràmetres tenen sobre les solucions (5.20) i sobre els valors òptims (5.21). Considerem en primer lloc l'efecte d'una variació a **b** sobre el valor òptim $F^* = F(x^*)$:

$$\frac{\partial F^*}{\partial \mathbf{b}^1} = \mathbf{c}^1(\mathbf{A}^{11})^{-1} = \mathbf{y}^{1^*}; \quad \frac{\partial F^*}{\partial \mathbf{b}^2} = \mathbf{0}$$

Així doncs,

$$\mathbf{y}^* = \frac{\partial F}{\partial \mathbf{b}}$$

De manera semblant, considerem ara l'efecte d'una variació de la constant de restricció en el problema dual:

$$\frac{\partial G^*}{\partial \mathbf{c}^1} = (\mathbf{A}^{11})^{-1}\mathbf{b}^1 = \mathbf{x}^{1^*}; \quad \frac{\partial G^*}{\partial \mathbf{c}^2} = \mathbf{0}$$

Així doncs,

$$\mathbf{x}^* = \frac{\partial G^*}{\partial \mathbf{c}}$$

En conseqüència, la sensibilitat del valor òptim de la funció objectiu a variacions en la constant de restricció està mesurada pel valor òptim de la variable dual corresponent. Aquesta interpretació és idèntica a la donada pel cas del problema de programació no lineal. Com vem fer notar allà, en alguns problemes económics d'assignació de recursos, les variables duals tenen una interpretació natural com "preus ombra".

El valor òptim d'una funció objectiu és independent d'una constant de restricció si la variable dual corresponent és zero. Aquest és un resultat raonable perque si la restricció no es operativa, variar una mica el valor de la constant de restricció no ha d'afectar al problema. De fet, a partir e (5.20):

$$\begin{aligned} \frac{\partial \mathbf{x}^{1^*}}{\partial \mathbf{b}^2} &= \mathbf{0}, \quad \frac{\partial \mathbf{x}^{2^*}}{\partial \mathbf{b}^2} = \mathbf{0} \\ \frac{\partial \mathbf{y}^{1^*}}{\partial \mathbf{c}^2} &= \mathbf{0}, \quad \frac{\partial \mathbf{y}^{2^*}}{\partial \mathbf{b}^2} = \mathbf{0} \end{aligned}$$

per tant variar el valor de la constrant de restricció en una restricció no operativa no afecta a la solució del problema. En problemes econòmics el fet de que la restricció no sigui operativa vol dir, per exemple, que la demanda és inferior a la oferta, condunt a un preu ombra zero. La solució aleshores, és independent de la oferta total disponible del be en qüestió donat que ja hi ha una oferta més que suficient en relació a l'us del be en l'òptim.

La sensibilitat de les solucions en respecte a constants de restricció quan aquestes són operatives també s’obté a partir de la diferenciació de (5.20):

$$\frac{\partial \mathbf{x}^{1^*}}{\partial \mathbf{b}^1} = (\mathbf{A}^{11})^{-1}$$
$$\frac{\partial \mathbf{y}^{1^*}}{\partial \mathbf{c}^1} = (\mathbf{A}^{11})^{-1}$$

de manera que els corresponents elements d’aquestes matrius són iguals.

La sensibilitat dels valors òptims de la funció objectiu a variacions en les constants de restricció són:

$$\frac{\partial F(\mathbf{x}^*)}{\partial \mathbf{c}^1} = (\mathbf{A}^{11})^{-1}\mathbf{b}^1 = \mathbf{x}^{1^*}$$
$$\frac{\partial G(\mathbf{y}^*)}{\partial \mathbf{b}^1} = \mathbf{c}^1(\mathbf{A}^{11})^{-1} = \mathbf{y}^{1^*}$$

Finalment, considerem els efectes de variacions en els coeficients de la matriu. Es clar que tots els elemente de la matriu $\mathbf{A}$ que no siguin els de la submatriu $m_1 \times n_1$ que denotem com $\mathbf{A}^{11}$ no tenen cap efecte sobre la solució ni sobre els valors òptims. Per la submatriu $\mathbf{A}^{11}$, diferenciant (5.21) obtenim,

$$\frac{\partial F(\mathbf{x}^*)}{\partial \mathbf{A}^{11}} = -(\mathbf{A}^{11})^{-1}\mathbf{b}^1\mathbf{c}^1(\mathbf{A}^{11})^{-1} = \frac{\partial G(\mathbf{y}^*)}{\partial \mathbf{A}^{11}}$$

on cada terme és una matriu $m_1 \times n_1$. Utilitzant (5.20):

$$\frac{\partial F(\mathbf{x}^*)}{\partial \mathbf{A}^{11}} = -\mathbf{x}^{1^*}\mathbf{y}^{1^*} = \frac{\partial G(\mathbf{y}^*)}{\partial \mathbf{A}^{11}},$$

de manera que

$$\frac{\partial F(\mathbf{x}^*)}{\partial a_{ji}} = -\mathbf{x}_j^*\mathbf{y}_i^* = \frac{\partial G(\mathbf{y}^*)}{\partial a_{ji}}, \quad i = 1, 2, \ldots, m_1; \quad j = 1, 2, \ldots, n_1.$$

5.4 L’algoritme del Simplex.

Els problemes de programació lineal es poden resoldre *geomètricament* si o be el número d’instruments o be el número de restriccions de desigualtat és dos o tres. En el cas de que el número d’instruments sigui dos o tres, la solució geomètrica és inmediata a partir de la representació del mapa de corbes de nivell, de la direcció de preferència, i del conjunt d’oportunitat. També si el número de restriccions de desigualtat és dos o tres, podem resoldre geomètricament de la forma indicada abans el problema dual, i podem utilitzar la solució del problema dual per trovar la solució del problema primal, donat que aleshores ja sabem el valor de la funció objectiu a la solució i les restriccions de desigualtat que es satisfan com igualtats.

L'*algoritme del simplex* és un mètode algebraic iteratiu per resoldre problemes de programació lineal amb un número qualsevol de restriccions de desigualtat i/o d'instruments. La idea general del seu funcionament consisteix en el següent: l'algoritme comena en un vèrtex qualsevol del conjunt d'oportunitat i a partir d'aquest es mou vers un vèrtex ve en la direcció que representi un increment en el valor de la funció objectiu. Aquest procediment iteratiu es va repetin fins que l'algoritme arriva a un vertex en el que no trova cap direcció en la que la funció objectiu augmenti el seu valor. Aquest vèrtex és la solució global del problema. Si totes les direccions al voltant d'aquest vèrtex impliquen disminucions en el valor de la funció objectiu, aleshores la solució és única; si hi ha alguna direcció en el que la funció no decreix, la solució no és unica i tots aquells vèrtices són solucions així com tots els punts intermedis. Donat que el número de vèrtices en el conjunt d'oportunitats és finit, aquest algoritme o be trova la solució al problema o be ens diu que la funció objectiu no es afitada després d'un número finit d'iteracions[1]

La millor manera de comprendre el metode del simplex consisteix en desenvolupar el seu funcionament en un exemple senzill. Considerem el problema de programació lineal següent:

$$
\begin{aligned}
\max_{x_1,x_2} F =& 3x_1 + 2x_2 \quad s.a. \\
& 2x_1 + x_2 6 \\
& x_1 + 2x_2 8 \\
& x_1 0, x_2 0
\end{aligned}
$$

No cal dir que aquest és un exemple extremadament senzill que es pot resoldre geomètricament, pero en qualsevol cas resultar útil per il.lustrar el funcionament de l'algoritme.

La dinàmica de l'algoritme per passar d'un vèrtex a un altre vèrtex ve es pot descomposar en un procediment en quatre etapes:

- El primer pas consisteix en introduir variables de folgana per transformar les restriccions de desigualtat en igualtats. Denominarem a aquestes variables

[1] Això el que vol dir és que per problemes en els que el número de vèrtex en el conjunt d'oportunitats és molt gran, l'algoritme no examina tots i cada un dels vèrtex. En aquest sentit es diu que l'algoritme del simplex és un algoritme de tipus restringit.

de folgana s_1 i s_2. Les restriccions del problema ara són

$$\begin{aligned}
&2x_1 + x_2 + s_1 = 6\\
&x_1 + 2x_2 + s_2 = 8\\
&x_1 0, x_2 0\\
&s_1 0, s_2 0
\end{aligned}$$

- El segon pas consisteix en localitzar una *solució factible*, és a dir un vèrtex del conjunt factible. Per simplicitat, i donat que en el nostre problema l'origen en E^2 és factible, prenem l'origen com la solució bàsica. En aquest punt obtenim els següents valors:

  $$x_1 = 0, \quad x_2 = 0, \quad s_1 = 6, \quad s_2 = 8, \quad F = 0.$$

  En aquest punt hem de distinguir entre les variables que a la solució bàsica són zero (i.e. $x_1, x_2,$) i les que no són zero (i.e. s_1 i s_2). Les primeres s'anomenen variables *no bàsiques*.

- El tercer pas és expresar la funció objectiu i les restriccions en termes de les variables no bàsiques, és a dir

  $$\begin{aligned}
  &F = 3x_1 + 2x_2\\
  &s_1 = 6 - 2x_1 - x_2\\
  &s_2 = 8 - x_1 - 2x_2
  \end{aligned} \tag{5.22}$$

- El quart pas és el moviment a un vèrtex ve. Per cada una de les varibles no bàsiques individualment considerades, determinem en quant es pot augmentar (sense violar les restriccions), i donat aquest increment, en quant augmentaria el valor de la funció objectiu. En les equacions (5.22) l'increment de qualsevol variable no bàsica està limitat per la no negativitat de les variables de folgana.

  Considerem x_1. Donat que d'acord amb la segona equació x_1 pot augmentar fins a tres, i d'acord amb la tercera equació pot augmentar fins a vuit, el màxim increment que satisfà ambdues restriccions és tres. El valor de la funció objectiu associat a aquest increment és nou. De forma paralela, el màxim increment per x_2 és 4 que dona lloc a un increment en el valor de la funció objectiu de vuit. Si ens movem al llarg de la direcció del més gran increment de valor de la funció objectiu, incrementem el valor de x_1 en tres, el que redueix s_1 a zero i s_2 a cinc. Així, s_1 substitueix a x_1 com variable no bàsica. (De igual forma el moviment podria haver estat $x_2 = 4$ de manera que $s_2 = 0$ donat que també representa un increment de la funció objectiu, no necessariament la direcció del més gran increment.)

Figura 5.5: La taula de l'exemple d'aplicació del mètode del simplex.

Figura 5.6: Transformació pivotal.

La nova solució bàsica factible és ara

$$x_1 = 3, \quad x_2 = 0, \quad s_1 = 0, \quad s_2 = 5, \quad F = 9.$$

on les variables no bàsiques són x_2 i s_1.

La transformació en aquest punt s'anomena *transformació pivotal* i representa la transformació computacional bàsica de l'algoritme del simplex. Fonamentalment aquesta transformació dona les noves variables bàsiques i la funció objectiu com funcions lineals de les variables no bàsiques.

En termes de la figura 5.5, la transformació pivotal involucra dos etapes quan es pivota sobre un element diferent de zero a_{ij}. La primera etapa és *normalitzar* dividint tots els elements de la fila pivotal i per l'element pivotal a_{ij}. La segona etapa és *eliminar* a base de restar multiples de la fila pivotal de les altres files per tal d'obtenir zeros per tots els elements de la columna pivotal j excepte per l'u de la posició pivotal. Pivotant sobre a_{ij} te doncs l'efecte de solucionar la i-èsima equació per x_j i utilitzar aquesta equació per eliminar aquesta variable de totes les altres equacions. En el nostre exemple, la taula original es mostra en la figura 5.5

Pivotant sobre l'element encerclat en la figura 5.5, obtenim la taula que es mostra en la figura 5.6 on els valors de les variables bàsiques apareixen a la darrera columna. (Ignorem, de moment que l'element a_{22} està encerclat.) En general, l'element pivotal ha de ser un element no zero a_{ij} (en el nostre exemple un 2) pel que el corresponent element en la darrera fila, c_j (en el nostre cas 3) és positiu i pel que la taula transformada no conté elements negatius en la darrera columna (en el nostre exemple els elements transformats són 3 i 5).

Continuant amb l'exemple, el següent pas és repetir el tercer pas, obtenir les

Figura 5.7: Segona transformació pivotal.

restriccions i la funció objectiu en termes de les variables no bàsiques com:

$$\begin{aligned} x_1 &= 3 - \frac{1}{2}x_2 - \frac{1}{2}s_1 \\ s_2 &= 8 - x_1 - 2x_2 = 5 - \frac{3}{2}x_2 + \frac{1}{2}s_1 \\ F &= 3x_1 + 2x_2 = 9 + \frac{1}{2}x_2 - \frac{3}{2}s_1 \end{aligned}$$

Aquests coeficients es podrien haver obtingut directament de les files de la taula transformada de la figura 5.6.

Repetint el quart pas, x_2 es pot augmentar fins a $\frac{10}{3}$ (i.e. fins el punt on s_2 esdevè negatiu), el que augmenta la funció F en $\frac{10}{6}$. Es clar que aquest és l'únic moviment factible donat que F es decreixent en s_1. Implementant la variació esmentada sobre x_2 obtenim una nova solució bàsica:

$$x_1 = \frac{4}{3}, \quad x_2 = \frac{10}{3}, \quad s_1 = 0, \quad s_2 = 0, \quad F = \frac{32}{3}. \tag{5.23}$$

on les variables no bàsiques són s_2 i s_1. En termes de la taula de la figura 5.6, pivotant sobre l'element encerclat obtenim la taula que es mostra en la figura 5.7.

Aquesta solució bàsica ens proporciona les equacions per repetir el tercer pas una vegada més. Les equacions són

$$\begin{aligned} x_1 &= \frac{4}{3} - \frac{2}{3}s_1 + \frac{1}{3}s_2 \\ x_2 &= \frac{10}{3} + \frac{1}{3}s_1 - \frac{2}{3}s_2 \\ F &= \frac{32}{3} - \frac{4}{3}s_1 - \frac{1}{3}s_2 \end{aligned} \tag{5.24}$$

que mostra que la solució bàsica (5.23) és de fet la solució al problema donat que la funció objectiu a (5.24) és decreixent tant en s_1 com en s_2. En termes de la figura 5.7, observem que tots els elements de la darrera fila (els coeficients de la funció objectiu) són zero o negatius. Aquesta taula de la figura 5.7 és òptima perque els elements de la darrera fila són no positius i els elements de la darrera columna són no negatius. Els elements de la darrera columna solucionen el problema primal. La solució obtinguda pel mètode del simplex és doncs:

$$x_1^* = \frac{4}{3}, \quad x_2^* = \frac{10}{3}, \quad F^* = \frac{32}{3}.$$

Els coeficients de les variables de folgana de la darrera equació corresponent a la funció objectiu (i.e. els elements diferents de zero de la darrera fila en la taula de la figura 5.7) són els negatius dels valors òptims de les variables duals. Això és així perque aquests coeficients representen les taxes a les que podria augmentar el valor de la funció objectiu si s_1 i s_2 poguessin ser negatius, o de forma equivalent, si b_1 i b_2 fossin més grans. Així, la solució del problema dual

$$\begin{aligned} \min_{y_1,y_2} G =& 6y_1 + 8y_2 \quad s.a. \\ & 2y_1 + y_2 3 \\ & y_1 + 2y_2 2 \\ & y_1 0, y_2 0 \end{aligned}$$

és:

$$y_1^* = \frac{4}{3}, \quad y_2^* = \frac{1}{3}, \quad G^* = \frac{32}{3}.$$

Els valors òptims de les variables duals mesuren la sensibilitat del valor òptim de la funció objectiu del problema primal F^* devant de variacions en les constants de les restriccions. Per exemple, si en la primera restricció el problema primal la constant fos 8 en lloc de 6, el valor òptim de la funció objectiu seria,

$$\Delta F^* = y_1^* \Delta b = \frac{4}{3}(8-6) = \frac{8}{3}$$

i el nou valor òptim seria

$$\frac{32}{3} + \frac{8}{3} = \frac{40}{3}.$$

La solució proporcionada per l'algoritme del simplex es mostra gràficament a la figura 5.8, on es mostra el moviment d'un vèrtex a un altre vèrtex ve fins que es trova la solució.

Tot aquest proces iteratiu es pot resumir de forma esquemàtica en la figura 5.9.

Figura 5.8: Il.lustració del mètode del simplex.

Figura 5.9: L’esquema de funcionament de l’algoritme del simplex.