

Ottimizzazione*

Come abbiamo spiegato nel capitolo precedente, l'idea di razionalità implica fra l'altro che un decisore cerca di individuare la migliore tra tutte le alternative a lui disponibili. In altri termini, egli cerca di *ottimizzare*. Per questa ragione il concetto di ottimizzazione è fondamentale in Microeconomia e problemi di ottimizzazione ricorrono continuamente in questo libro. Sebbene il particolare contesto economico dei vari problemi possa essere diverso — la scelta del consumatore di un paniere di beni, la decisione produttiva di un'impresa, la scelta di allocazione delle risorse da parte di un pianificatore a livello di intero sistema economico —, tutti tendono ad avere una struttura di base comune. È quindi utile e istruttivo investire un po' di tempo nel considerare problemi generali di ottimizzazione, astraendo da qualsiasi contesto particolare.

Il lettore con qualche base di matematica e gusto per l'astrazione potrà trovare utile leggere questo capitolo prima di procedere ai modelli economici dei capitoli successivi. In tal modo diverrà consapevole della natura di *problemi di ottimo* a mano a mano che li incontrerà. Una strategia altrettanto valida, comunque, è quella di affrontare questo capitolo dopo aver studiato i modelli di cui ai Capitoli 3, 7, 8, 9, 15 e 17, quando si avverta l'esigenza di una discussione generale sulla loro struttura comune.

La comprensione del materiale di questo capitolo non è strettamente *necessaria* per lo studio dei modelli economici che vengono presentati successivamente. Il lettore che li studia, tuttavia, può notare le similitudini sottostanti alla loro struttura. Può pure essere incuriosito dal fatto che quando si registrano mutamenti piuttosto ragionevoli nel modo in cui le curve sono tracciate, ne seguono effetti apparentemente strani nella natura dei risultati. Per esempio, non è irragionevole supporre che le curve d'indifferenza del consumatore siano in ogni punto del quadrante positivo più ripide della sua linea di bilancio — nulla negli assiomi esposti nel Capitolo 3 impedisce questa possibilità. Ma cosa accade alla desiderata "soluzione di tangenza" per l'equilibrio del consumatore, e alla "condizione necessaria" che il tasso marginale di sostituzione *uguagli* il rapporto tra i prezzi? Queste situazioni possono essere soddisfatte, ma al costo di permettere consumo negativo di uno dei beni (provate a tracciare il diagramma). Ha senso tutto ciò?

Un esame della teoria generale dei problemi di ottimizzazione fornisce un quadro di riferimento per i tipici modelli economici. Ci rende consapevoli delle ipotesi che sono state tacitamente introdotte nell'analisi al fine di assicurare che i risultati possiedano certe proprietà e ci permette di riformulare le risposte quando tali ipotesi implicite siano rimosse. Come minimo, tale esame accresce notevolmente la nostra consapevolezza della natura e del significato degli stessi modelli economici.

1. La struttura dei problemi di ottimizzazione

Finora abbiamo parlato di ottimizzazione come l'atto di scegliere la "migliore" alternativa tra quelle disponibili. Si tratta della descrizione di come le decisioni (ossia le scelte tra alternative) sono o dovrebbero essere prese. Ora procediamo oltre questa semplice nozione ed esaminiamo in dettaglio i quesiti che si possono porre circa i problemi di ottimizzazione e i concetti che sono stati sviluppati per dare loro una risposta.

Tutti i problemi di ottimizzazione consistono di tre elementi:

Le variabili decisionali

Sono queste le variabili i cui valori ottimali sono da determinarsi. Per esempio:

1. Un'impresa desidera conoscere il livello del proprio prodotto o output che genera il massimo profitto. L'output è la variabile decisionale.
2. Un'impresa vuole conoscere le quantità di lavoro, tempo-macchina e materie prime da impiegare così da produrre un dato livello di output al minimo costo. Variabili decisionali sono lavoro, tempo-macchina e materie prime.
3. Un consumatore desidera quel paniere di beni che gli sia possibile acquistare dato il proprio reddito e che gli procuri la massima utilità. Qui le variabili decisionali sono le quantità di beni.

In Economia quasi sempre si suppone che la quantità di ogni variabile decisionale sia misurabile con un numero reale. Vi può essere qualsiasi numero finito di variabili decisionali in un problema particolare. Noi ci troveremo normalmente a discutere problemi a una o due variabili, giacché problemi con un maggior numero di variabili non possono essere immediatamente rappresentati in un diagramma bidimensionale. La perdita di generalità è di solito più apparente che reale.

La funzione obiettivo

Questa funzione fornisce la specificazione matematica della relazione fra le variabili decisionali da un lato e certe variabili il cui valore desideriamo *massimiz-*

zare c
obietti

1. il p
 2. il c
 3. un
- può

Nei cas
da min

Il let
la terza
ze mon
invece,
mo sog
tazione
dettagli
cludere
ne obie
obiettivi
mediata
zioni n
massimi
mini deg
pagare n
prio den
ticolare,
decisione
di tale de
a meno c
to delle p
parte, pi
sforzo fa
tale aspet
di una va
un qualch

Nel seg
che grand
bili decis
un'utile in
me una fu
di prefer

zare o *minimizzare* dall'altro. Perciò nei tre esempi visti sopra, la funzione obiettivo metterebbe in relazione:

1. il profitto con il livello di output;
2. il costo agli ammontari di lavoro, tempo-macchina e materie prime;
3. un indice della soddisfazione del consumatore e le quantità dei beni che egli può comprare.

Nei casi (1) e (3) le funzioni obiettivo saranno da massimizzare, e nel caso (2) da minimizzare, rispetto alle variabili decisionali rilevanti.

Il lettore — ci auguriamo — avrà percepito la differenza che intercorre fra la terza di tali funzioni obiettivo e le prime due. Profitto e costo sono grandezze monetarie, la cui misurazione non presenta problemi. Che cosa intendiamo, invece, per “un indice di soddisfazione”? La soddisfazione è uno stato d'animo soggettivo e non è immediatamente ovvio che sia suscettibile di rappresentazione mediante un indice numerico. Nel prossimo capitolo considereremo dettagliatamente una serie di ipotesi che, se verificate, ci permetteranno di concludere come per il problema del consumatore sia *possibile* definire una funzione obiettivo numerica che egli massimizza. In fin dei conti tutte le funzioni obiettivo, perfino quelle che sembrano a prima vista coinvolgere quantità immediatamente misurabili come profitti e costi, altro non sono che rappresentazioni numeriche di ordinamenti di preferenze. Un imprenditore cerca di massimizzare il profitto perché *preferisce* più profitto a meno profitto, in termini degli usi che ne può fare. Cerca di minimizzare i costi perché *preferisce* pagare meno che pagare di più, date tutte le altre possibilità di spendere il proprio denaro. Queste due funzioni obiettivo in realtà rappresentano un caso particolare, in cui il decisore è interessato a un solo esito misurabile della propria decisione (profitto o costo), e il suo ordinamento di preferenze relative ai valori di tale decisione è di forma particolarmente semplice: egli preferisce sempre più a meno o meno a più. In una simile situazione possiamo ignorare l'ordinamento delle preferenze e operare direttamente in termini del singolo esito. Se d'altra parte, più di un esito della decisione fosse rilevante (per esempio, profitto più sforzo fatto dall'imprenditore), allora non potremmo ignorare il più fondamentale aspetto dell'ordinamento delle preferenze. Un'alternativa può implicare più di una variabile e meno dell'altra rispetto a una seconda alternativa, cosicché un qualche tipo di valutazione è inevitabile.

Nel seguito di questo capitolo supporremo semplicemente di massimizzare qualche grandezza che sia un numero reale e che sia una data funzione delle variabili decisionali del problema. La discussione appena fatta, tuttavia, ci fornisce un'utile interpretazione della funzione obiettivo: possiamo cioè considerarla come una funzione che assume valori diversi delle variabili decisionali in ordine di preferenza cosicché possiamo individuare quelli preferiti.

L'insieme ammissibile

Abbiamo finora parlato liberamente di "alternative disponibili". Una parte essenziale di ogni problema di ottimizzazione è l'esatta specificazione di quali alternative sono disponibili per il decisore. L'insieme di alternative disponibili è definito "insieme ammissibile". Dato che le alternative sono normalmente considerate "punti" nello spazio, gli insiemi ammissibili costituiscono ciò che in matematica generalmente definiamo come "insiemi di punti".

Vi sono tre modi in cui l'insieme ammissibile può essere specificato:

- mediante enumerazione diretta, per esempio attraverso l'affermazione: le alternative sono A, B, C, \dots . Chiaramente se l'insieme ammissibile comprende una sola alternativa, o nessuna, il problema di ottimizzazione è banale.
- mediante una o più disuguaglianze che definiscano *direttamente* l'insieme di valori alternativi della o delle variabili decisionali.
- mediante una o più *funzioni* o equazioni che definiscano un insieme di valori alternativi.

Esempi di queste due ultime modalità possono essere forniti sulla base dei problemi menzionati precedentemente. Quindi, nel problema (1) escludiamo livelli negativi di output ma ci aspetteremo qualsiasi output positivo come possibile. Perciò diciamo che l'output deve essere maggiore o uguale a zero, cioè $y \geq 0$ dove y denota l'output. L'insieme ammissibile è qui direttamente definito da una *disuguaglianza debole*.

Il secondo problema ci dice che possono essere considerate solo quelle combinazioni di input che forniscano il desiderato livello di output. In tal caso l'insieme ammissibile è definito mediante una funzione. Quindi supponiamo di avere la *funzione di produzione*: $y = f(L, M, R)$ dove y è l'output, L, M, R , sono lavoro, tempo-macchina e materie prime rispettivamente. Sia ora y^0 il livello di output richiesto. Allora l'equazione:

$$y^0 = f(L, M, R) \quad (1.1)$$

definisce l'insieme di valori di L, M, R che sono ammissibili. Si noti inoltre che l'equazione (1.1) potrebbe essere soddisfatta per valori negativi di una o più variabili decisionali, il che implica che tali valori negativi sembrerebbero legittimi. Ovviamente questo non è il caso, dato che non consideriamo sensati valori negativi per tali variabili, per cui vogliamo escludere una simile evenienza. Aggiungiamo allora i vincoli:

$$L \geq 0 \quad R \geq 0 \quad M \geq 0 \quad (1.2)$$

che, insieme con la (1.1), definiscono l'insieme ammissibile.

Nell'esempio del problema del consumatore possiamo anzitutto affermare che è impossibile che egli consumi quantità negative di beni. Quindi, rappresentan-

do con
biamo

$$x_1 \geq$$

Ma v
ma. Og
 x_2, \dots, x_n
certa co

$$p_1 x_1 +$$

Il con
eccedere
suguaglia

$$p_1 x_1 +$$

unitamer

In sost
problema
alternativ
per restri
zione nell
Se il prob
tero spazi
sionali) e
l'insieme

Per rica
cisionali,
nello scegli
stra teoria
di trovare
decisionali
to un sino

Esercizi

1.* Si desc
per i s
1. Uno
il tre

do con x_1, x_2, \dots, x_n le quantità di beni che il consumatore può acquistare, abbiamo immediatamente le n disuguaglianze:

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad \dots \quad x_n \geq 0 \quad (1.3)$$

Ma vi è un'altra limitazione per l'insieme ammissibile implicito nel problema. Ogni bene ha un prezzo. Siano p_1, p_2, \dots, p_n i prezzi delle quantità x_1, x_2, \dots, x_n rispettivamente. Allora la spesa complessiva del consumatore per una certa combinazione di quantità dei beni è:

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n = \sum_{i=1}^n p_i x_i \quad (1.4)$$

Il consumatore ha un certo reddito, M , e la sua spesa di consumo non può eccedere tale livello. Perciò nel problema l'insieme ammissibile è dato dalla disuguaglianza:

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n \leq M \quad (1.5)$$

unitamente alle disuguaglianze (1.4).

In sostanza, in ognuno dei problemi portati a esempio, come in qualunque problema di ottimizzazione, è necessario definire in maniera esatta l'insieme delle alternative disponibili. Le funzioni e le disuguaglianze riportate sopra e usate per restringere il numero di alternative che possono essere prese in considerazione nella definizione di insieme ammissibile prendono il nome di "vincoli". Se il problema non possiede vincoli allora l'insieme ammissibile consiste nell'intero spazio n -dimensionale dei vettori reali (dove n è il numero di variabili decisionali) e il problema è detto "non vincolato"; la presenza di vincoli restringe l'insieme ammissibile a un sottoinsieme dell'intero spazio.

Per ricapitolare: un problema di ottimizzazione è costituito da variabili decisionali, una funzione obiettivo e un insieme ammissibile. Il problema consiste nello scegliere l'alternativa preferita all'interno dell'insieme ammissibile, e la nostra teoria ci permette generalmente di rappresentare tutto ciò come il problema di trovare il massimo o il minimo della funzione obiettivo rispetto alle variabili decisionali, soggetto ai vincoli. Per questa ragione, ottimizzazione è considerato un sinonimo di massimizzazione o minimizzazione vincolata.

Esercizi

- 1.* Si descrivano variabili decisionali, funzione obiettivo e insieme ammissibile per i seguenti problemi di ottimizzazione:
1. Uno studente può tornare da scuola a casa, a piedi, con l'autobus o con il treno. Conosce con esattezza quanto tempo gli ci vuole in ciascun modo

e quanto gli costerà. Il problema è quello di scegliere il "miglior" modo per tornare a casa.

2. Una persona intende iniziare una dieta che generi il minor numero possibile di calorie, soggetto a un certo minimo, e che assicuri il consumo di certe quantità minime di vitamine. La persona, inoltre, non può eccedere il proprio bilancio settimanale di spesa per cibo. Conosce prezzo, numero di calorie e contenuto di vitamine di ogni singola unità di cibo. In aggiunta esistono delle costose pillole ricche di vitamine e prive di calorie. Come determinare la dieta "ottimale"?
 3. Una massaia ha la scelta di fare la spesa al mercato A che si trova vicinissimo a casa, o di prendere l'autobus e fare acquisti al mercato B dove i prezzi sono relativamente più bassi. Il problema è di scegliere se fare la spesa al mercato A, B, oppure a tutti e due.
- 2.* Si traccino i grafici degli insiemi ammissibili definiti dai seguenti vincoli:

$$(i) \quad x_2 + 2x_1 \leq 4$$

$$(ii) \quad x_2 + 3x_1 < 6$$

$$(iii) \quad x_2 + 2x_1 \leq 4$$

$$x_2 + 4x_1 \leq 6$$

$$x_1 \geq 0 \quad x_2 \geq 0$$

$$(iv) \quad x_2 + 2x_1 = 4$$

$$x_2 + 3x_1 = 7$$

$$(v) \quad x_2 + 2x_1 = 4$$

$$x_2 + 3x_1 = 7$$

$$x_1 \geq 0 \quad x_2 \geq 0$$

$$(vi) \quad x_2 - 2x_1^2 \leq 0$$

$$x_2 \geq 0 \quad x_1 \geq 0$$

$$(vii) \quad x_2 + x_1^2 = 4$$

$$(viii) \quad x_2 + x_1^2 = 4$$

$$x_1 \geq 0 \quad x_2 \geq 0$$

$$(ix) \quad x_2 + x_1^2 \leq 4$$

$$x_2 + 3x_1 \leq 6$$

$$x_1 \geq 0 \quad x_2 \geq 0$$

$$(x) \quad x_2 + x_1^2 = 4$$

$$x_2 + 3x_1 = 6$$

2. Soluzione: quesiti e concetti

La *soluzione* di un problema di ottimizzazione è costituita da quel vettore di valori delle variabili decisionali che appartiene all'insieme ammissibile e che produce il valore massimo o minimo della funzione obiettivo per quell'insieme ammissibile. È opportuno a questo punto introdurre un poco di notazione. Rappresentiamo la funzione obiettivo come:

$$f(x_1, x_2, \dots, x_n) = f(x) \quad (2.1)$$

dove x è il vettore di n elementi delle variabili decisionali. Per comodità supporremo che il problema sia sempre quello di *massimizzare* $f^{(1)(*)}$. Denotiamo l'in-

(*) I numeri soprascritti (in parentesi) si riferiscono alle note al termine di questo capitolo.

sieme a
vettore

$f(x^*)$

che è un
l'insieme
il vettore
Vi son
con rigua

Esistenza

Come es
una solu
per supp
nomica s
timizzazi
basi per l

Soluzioni

Una solu
funzione
all'interno
soluzione

$f(x^{**})$

dove N^{**}
differenza
re diventa
guaglianze
corrispond
come N^*
(2.3), ma s
tre x^{**} no

Il proble
mo per tro
li. Ci intere
locale è pur
aver a che
una funzio
gior dettagl

sieme ammissibile dei vettori x con S . Allora una soluzione del problema è un vettore di variabili decisionali, x^* , avente la proprietà:

$$f(x^*) \geq f(x) \quad \text{per ogni } x \in S, \quad x^* \in S \quad (2.2)$$

che è un altro modo di esprimere il fatto che x^* massimizza f relativamente all'insieme S . Per definizione del problema, il nostro scopo è quello di trovare il vettore x^* .

Vi sono alcuni importanti quesiti di carattere generale che ci possiamo porre con riguardo alla soluzione di un problema di ottimizzazione:

Esistenza

Come essere sicuri, prima di cercare di risolvere un problema particolare, che una soluzione effettivamente esiste per esso? Dopotutto non abbiamo ragioni per supporre che ogni problema *debba* possedere una soluzione. In teoria economica spendiamo molto tempo nell'analizzare le soluzioni dei problemi di ottimizzazione. Dobbiamo perciò preoccuparci che le nostre teorie possiedano le basi per la propria esistenza, altrimenti l'analisi sarà internamente inconsistente.

Soluzioni locali e globali

Una soluzione *globale* è una soluzione che soddisfi la (2.2): in quel punto la funzione obiettivo assume un valore che non è inferiore ad alcun altro punto all'interno dell'insieme ammissibile. È quindi la soluzione che ricerchiamo. Una soluzione locale, invece, soddisfa la condizione:

$$f(x^{**}) \geq f(x) \quad \text{per ogni } x \in N^{**} \subset S \quad (2.3)$$

dove N^{**} è un insieme di punti in un *intorno* di x^{**} . La figura 1 illustra la differenza. In essa ipotizziamo un'unica variabile decisionale, cosicché il vettore diventa lo scalare x . L'insieme ammissibile è definito solamente dalle disuguaglianze: $x \geq 0$, $x \leq x^0$. La funzione obiettivo $f(x)$ ha due picchi, uno in corrispondenza di x^* e l'altro di x^{**} . Gli intorni di questi punti sono definiti come N^* e N^{**} rispettivamente. Chiaramente entrambi i punti soddisfano la (2.3), ma solamente x^* soddisfa la (2.2). Perciò x^* è un massimo assoluto mentre x^{**} non lo è.

Il problema principale è costituito dal fatto che *tutti i metodi di cui disponiamo per trovare delle soluzioni permettono di individuare solo dei massimi locali*. Ci interessa quindi chiederci quali siano le condizioni per cui ogni massimo locale è pure un massimo globale. Dalla figura 1 si deduce che la risposta deve aver a che fare con la forma della funzione obiettivo (si pensi per esempio ad una funzione che abbia un solo picco): esploreremo questa questione con maggior dettaglio più avanti.

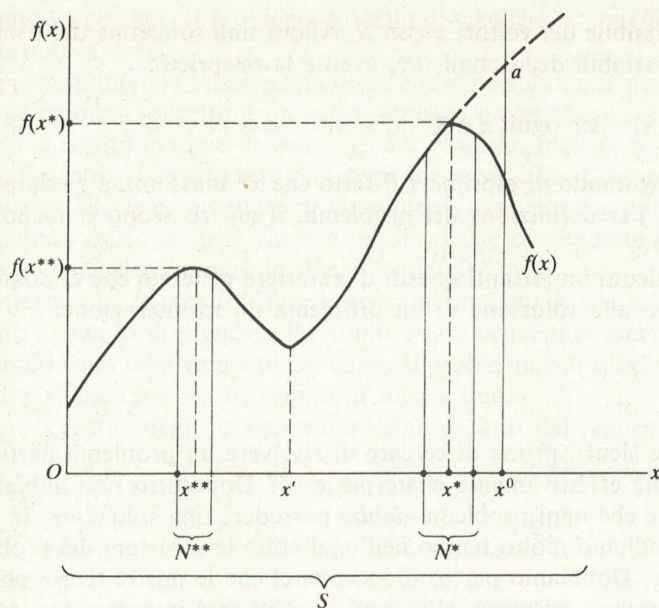


Fig. 1

Unicità

È concepibile che esista più di un massimo assoluto (per esempio il primo picco in figura 1 è alto quanto il secondo, ovvero la funzione ha un segmento orizzontale per l'insieme N^* di valori di x). Gli economisti generalmente ipotizzano soluzioni uniche, per cui è importante considerare le condizioni sotto le quali tale fatto si verifica.

Soluzioni interne

Anzitutto diamo per nota la distinzione tra punto interno e di confine di un insieme. Quindi, nella figura 1, i punti $x=0$ ed $x=x^0$ sono punti di confine, mentre tutti gli altri sono punti interni. Allora una *soluzione interna* è un punto interno che soddisfi la (2.2), mentre una *soluzione di confine* è un punto di confine che soddisfi quella condizione. Nella figura x^* è una soluzione interna. Se però la funzione f avesse la forma indicata dalla linea tratteggiata nel diagramma, allora vi sarebbe una soluzione di confine in corrispondenza di x^0 . L'importanza di questa distinzione risiede nelle conseguenze derivanti da una variazione in un vincolo (che in generale cambia la posizione di un confine dell'insieme ammissibile: si veda la Domanda 2 degli Esercizi). In seguito a una

piccola influenza di confine in figura 1, la soluzione portante dei vincoli (cioè il vincolo attivo) è il vincolo di confine. È

Possibile operare su una parte di un vincolo vi è una porzione di vincolo, 1, quando mentre il vincolo è non vincente, mentre più

Individuo

Dato che un vincolo di un vincolo e dei vincoli, no di ottimalità un contesto,ificando il problema, caratteristiche

Tale definizione prima enunciata di ottimo (e di ottimo sono), e quindi, ulteriore informazione a quelle di condizioni

A scopo di si suppone x . Allora, il valore per la funzione esibita, vata prima ma anche vincente. Si nota la derivata

piccola variazione in un vincolo una soluzione interna non verrà probabilmente influenzata: il punto di ottimo resterà invariato. D'altra parte, una soluzione di confine potrebbe benissimo essere influenzata: per esempio, nel caso in cui in figura 1 il punto x^0 fosse ottimale, uno spostamento dei confini cambierebbe la soluzione. Gran parte della Microeconomia si preoccupa di predire comportamenti desunti dall'analisi di variazioni nelle soluzioni ottimali al variare dei vincoli (si veda l'analisi del consumatore nel prossimo capitolo, per esempio). È quindi importante sapere se una soluzione è interna o di confine.

Possiamo pure formulare questa questione in termini di vincoli *operanti* e *non-operanti*. Un vincolo è operante se vi è una soluzione di confine che giace sulla parte di confine definita da quel vincolo. Un vincolo è perciò non-operante se vi è una soluzione interna, ovvero una soluzione di confine che giaccia sulla porzione di confine definita da un altro vincolo. A titolo di esempio, in figura 1, quando la soluzione è x^* , entrambi i vincoli $x \geq 0$ e $x \leq x^0$ sono non-operanti mentre nel caso in cui la soluzione corrisponde a x^0 , solamente il primo vincolo è non-operante. Possiamo sempre determinare una variazione sufficientemente piccola in un vincolo non-operante tale da lasciare la soluzione inalterata.

Individuazione della soluzione

Dato che una soluzione esiste, desideriamo individuarla. Nella caso di soluzione di un problema concreto otterremo valori numerici della funzione obiettivo e dei vincoli, e cercheremo poi di formulare procedure di calcolo che permettano di ottenere una soluzione quanto prima e più economicamente possibile. In un contesto teorico, tuttavia, lavoriamo solo con funzioni generiche, spesso specificando poco più dei segni delle derivate prime e seconde. Di conseguenza, il problema non è in effetti *calcolare* delle soluzioni, quanto *descrivere* le loro caratteristiche essenziali nel modo analiticamente più conveniente.

Tale descrizione è fatta in termini di *condizioni necessarie e sufficienti*: dapprima enunciamo certe condizioni che devono essere soddisfatte da ogni punto di ottimo (sebbene possano essere anche soddisfatte da punti che di ottimo non sono), e queste rappresentano le condizioni necessarie; quindi aggiungiamo un'ulteriore insieme di condizioni tali per cui ogni punto che le soddisfa unitamente a quelle necessarie deve essere un punto di ottimo. Perciò, questi due gruppi di condizioni insieme sono sia necessarie che sufficienti.

A scopo illustrativo, si prenda il caso di una massimizzazione non vincolata e si supponga in figura 1 che non esista alcun vincolo sui valori ammissibili di x . Allora, al valore x^* , la derivata $f'(x^*)=0$, e si può dimostrare che ciò deve valere per ogni massimo locale. Tuttavia, è anche vero che $f'(x')=0$, ma la funzione esibisce un minimo locale in quel punto. Quindi la condizione che la derivata prima sia uguale a zero è soddisfatta in tutti i punti di massimo locale, ma anche in altri punti, cosicché costituisce condizione necessaria ma non sufficiente. Si noti inoltre che a mano a mano che x cresce passando attraverso x^* , la derivata $f'(x^*)$ passa da valori positivi a valori negativi, cioè è decrescente,

e ciò vale solo per massimi locali. Ne segue che condizione necessaria e sufficiente per un massimo locale è che $f'(x^*)=0$ e $f''(x^*)<0$, poiché queste sono soddisfatte simultaneamente da *tutti* i massimi locali e *solo* per essi⁽²⁾. Si noti che le condizioni sono dette necessarie e sufficienti solo per massimi locali, giacché, come evidenziato dal grafico, esse non sono in grado di discriminare tra i punti x^* e x^{**} . Da cui segue la nostra discussione precedente sugli ottimi assoluti e relativi come un problema a sé. Descrivere un punto di ottimo in termini di condizioni necessarie e sufficienti significa "individuarlo" in termini delle sue caratteristiche generali anziché del suo valore numerico. Forse in modo sorprendente, siamo spesso in grado di dire parecchie cose sulla semplice base di una simile descrizione in termini generali.

Nella discussione di queste questioni concernenti le soluzioni di problemi di ottimizzazione facciamo uso di proprietà molto generali per funzioni obiettivo e insieme ammissibili. Elencheremo ora tali proprietà e procederemo poi a rispondere ai quesiti sollevati finora.

Continuità della funzione obiettivo

Una funzione $y=f(x)$ è continua se non vi sono interruzioni nel suo grafico o, rozzamente, se questo può essere tracciato senza sollevare la penna dal foglio. Le funzioni in (b) e (c) in figura 2 non sono continue, mentre quella in (a) lo è. In (b) $f(x)$ diventa arbitrariamente grande per $x=x^0$ (tende a infinito) e in (c) $f(x)$ salta da y^1 a y^2 in corrispondenza di x^0 . Quando vi sia più di una variabile nella funzione obiettivo, il concetto di continuità rimane valido: non vi dovrebbero essere salti o interruzioni nel grafico della funzione, ma diventa più difficile (impossibile nel caso di più di tre variabili) illustrarne graficamente l'effetto.

Concavità della funzione obiettivo

Nella figura 3 si trovano i grafici di quattro tipi di funzione. Una funzione con la curvatura rappresentata in (a) è detta *concava*, in (c) è detta *convessa*, men-

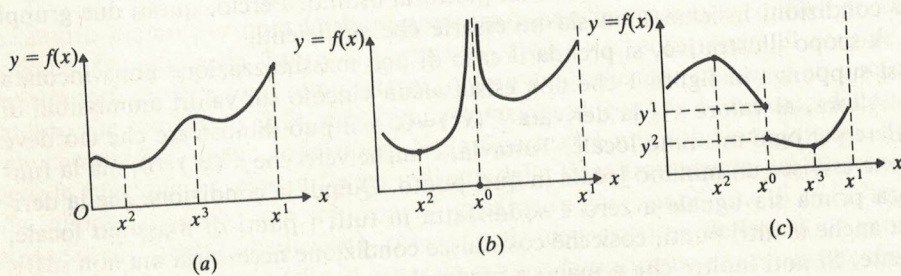


Fig. 2

$\frac{1}{2}f(x^0) + \frac{1}{2}f(x^1)$

tre in (b).
Se volessi
va, una p
che a mar
scenti di
e poi neg
zione, f'
funzione
 $f''(x) > 0$
sto appro
più punti
la definiz
Per ovv
più gener
se prendi
funzione
zione), al
sopra dell
te sotto la
con la lin
la linea.

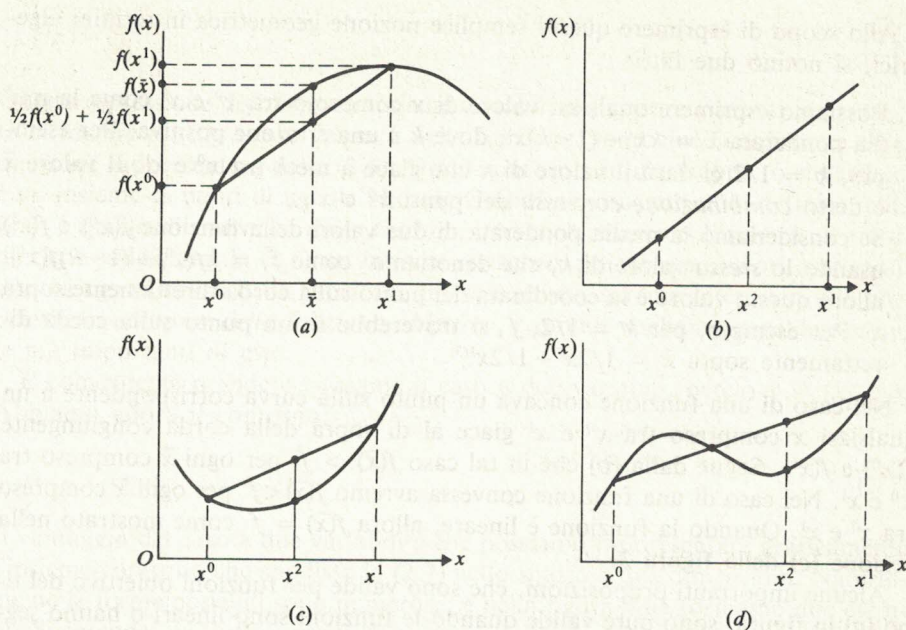


Fig. 3

tre in (b) è chiaramente lineare e infine in (d) non è né concava né convessa. Se volessimo trovare un modo non geometrico per definire una funzione concava, una possibilità sarebbe di farlo in termini di derivata seconda $f''(x)$. Si noti che a mano a mano che tracciamo delle tangenti alla curva in (a) per valori crescenti di x , tali tangenti vengono ad avere pendenze positive sempre più piccole e poi negative sempre più grandi, suggerendo che la derivata prima della funzione, $f'(x)$, decresce. Di conseguenza, possiamo esprimere la concavità di una funzione con la condizione $f''(x) < 0$. Analogamente, convessità è definita da $f''(x) > 0$ (si fornisca il ragionamento sottostante). Vi è però un problema in questo approccio: una funzione può non essere differenziabile in un punto (o in più punti), perché possiede un angolo in corrispondenza di esso (essi), e quindi la definizione non è utilizzabile (si costruisca un esempio).

Per ovviare a tale difficoltà definiamo concavità e convessità in termini di una più generale proprietà che la funzione possiede. Si noti che nella (a) in figura, se prendiamo due punti qualsiasi come x^0 e x^1 e congiungiamo i valori della funzione corrispondenti $f(x^0)$ e $f(x^1)$ con una linea retta (una *corda* della funzione), allora il grafico della funzione tra i due valori giace dappertutto al di sopra della linea. Per una funzione convessa, il grafico di essa giace interamente sotto la linea. Nel caso lineare ovviamente, il grafico della funzione coincide con la linea, mentre nel caso (d) della figura il grafico passerà da sopra a sotto la linea.

Allo scopo di esprimere questa semplice nozione geometrica in termini algebrici, si notino due fatti:

1. Possiamo esprimere qualsiasi valore di x compreso tra x^0 e x^1 come la media ponderata $\bar{x} = kx^0 + (1-k)x^1$, dove k è una frazione positiva. Per esempio, $k = 1/2$ ci darà il valore di x che giace a metà tra x^0 e x^1 . Il valore \bar{x} è detto *combinazione convessa* dei punti x^0 e x^1 .
2. Se consideriamo la media ponderata di due valori della funzione $f(x^0)$ e $f(x^1)$ usando lo stesso valore di k , che denotiamo come $\bar{f} = kf(x^0) + (1-k)f(x^1)$, allora questo valore è la coordinata del punto sulla corda direttamente sopra x . Per esempio, per $k = 1/2$, \bar{f} si troverebbe in un punto sulla corda direttamente sopra $\bar{x} = 1/2x^0 + 1/2x^1$.

Nel caso di una funzione concava un punto sulla curva corrispondente a un qualsiasi x compreso tra x^0 e x^1 giace al di sopra della corda congiungente $f(x^0)$ e $f(x^1)$. Segue dalla (b) che in tal caso $f(\bar{x}) > \bar{f}$ per ogni \bar{x} compreso tra x^0 e x^1 . Nel caso di una funzione convessa avremo $f(\bar{x}) < \bar{f}$ per ogni \bar{x} compreso tra x^0 e x^1 . Quando la funzione è lineare, allora $f(\bar{x}) = \bar{f}$ come mostrato nella sezione (c) della figura 3.

Alcune importanti proposizioni, che sono valide per funzioni obiettivo del tipo (a) in figura, sono pure valide quando le funzioni sono lineari o hanno segmenti lineari. È perciò utile definire *funzioni concave* quelle funzioni aventi la proprietà:

$$f(\bar{x}) \geq \bar{f} \quad (2.4)$$

dove $\bar{x} = kx^0 + (1-k)x^1$ e $\bar{f} = kf(x^0) + (1-k)f(x^1)$, per $0 \leq k \leq 1$, cosicché anche funzioni lineari o aventi segmenti lineari rientrano nella classe delle funzioni concave. Una funzione che soddisfi la (2.4) come disuguaglianza *forte* è allora detta *strettamente concava* e questo è il termine che dovrebbe essere usato per la funzione in figura 3(a). Allo stesso modo una *funzione convessa* soddisfa la condizione:

$$f(\bar{x}) \leq \bar{f} \quad (2.5)$$

e una funzione *strettamente convessa* soddisfa la condizione a disuguaglianza stretta. Si noti che (2.4) e (2.5) considerate insieme implicano che una funzione lineare è *sia* concava *che* convessa, sebbene non sia né strettamente concava né strettamente convessa.

Funzioni quasi-concave

Data una funzione $y = f(x_1, x_2, \dots, x_n)$ possiamo scegliere un numero c e definire:

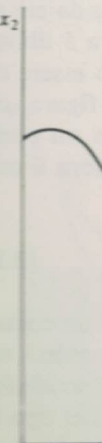
$$f(x) = c$$

Questa espressione di fornire una equazione razionale rappresenta un insieme (2.6) è quel di x che soddisfano a uguaglianza le più importanti dei c

È conveniente abbiamo a

$$f(x_1, x_2)$$

Il vantaggio l'insieme-concavamente le funzioni. Si intendono soddisfare a un dato intorno di continuità può salti nel gra



$$f(x) = c \quad (2.6)$$

Questa espressione definisce un insieme di valori del vettore x aventi la proprietà di fornire tutti lo stesso valore c della funzione: sono cioè la soluzione all'equazione rappresentata dalla (2.6). Ora il contorno di una mappa (o superficie) è un insieme di punti di uguale altezza. Quindi un concetto utile definito dalla (2.6) è quello di *contorno della funzione* $f(x)$ e chiameremo l'insieme di valori di x che soddisfano la (2.6) un *insieme-contorno*, poiché quei valori corrispondono a uguali valori della funzione. In varie istanze saremo interessati alle *proprietà dei contorni della funzione obiettivo*, e quindi vogliamo considerare qui le più importanti di esse.

È conveniente prendere in esame il caso a due variabili, perciò $x = (x_1, x_2)$. Abbiamo allora il contorno:

$$f(x_1, x_2) = c \quad (2.7)$$

Il vantaggio del caso a due variabili è che possiamo rappresentare graficamente l'insieme-contorno che soddisfa la (2.7) nello spazio a due dimensioni. Naturalmente le forme possibili sono numerose, e nella figura 4 ne forniamo due esempi. Si intende che tutti i punti di una linea di contorno c sono vettori x che soddisfano un'equazione per un contorno del tipo (2.7) e quindi appartengono a un dato insieme-contorno. Il diagramma illustra un'importante proprietà del contorno di una funzione, e cioè quello di continuità. Come in precedenza, continuità può essere concepita intuitivamente come assenza di interruzioni, buchi, salti nel grafico. La continuità di una funzione e del suo contorno sono stretta-

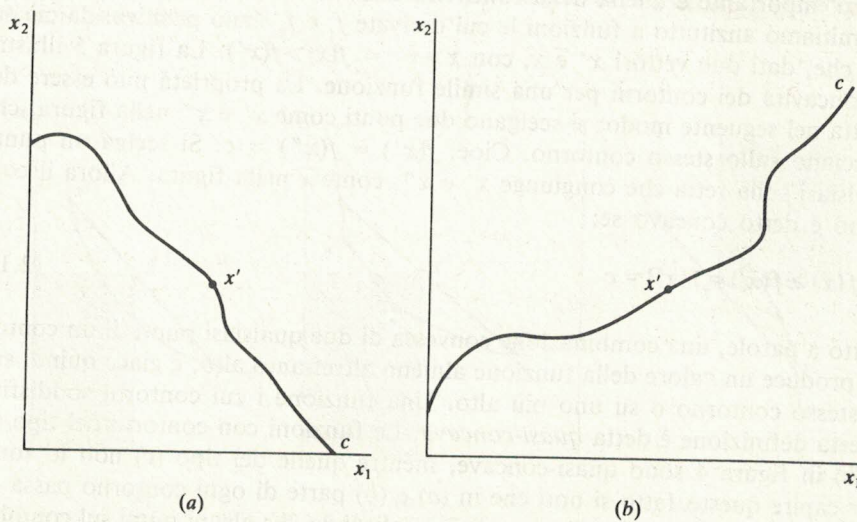


Fig. 4

mente collegati: può essere dimostrato che continuità di una funzione implica continuità dei suoi contorni.

Al fine di esplorare ulteriormente le proprietà dei contorni, procediamo oltre il concetto di continuità e supponiamo che la funzione $f(x)$ sia differenziabile. Allora, differenziando la (2.7) totalmente si ottiene:

$$\begin{aligned} df &= f_1 dx_1 + f_2 dx_2 \\ &= 0 \text{ poiché } c \text{ è una costante.} \end{aligned} \quad (2.8)$$

Si noti che i differenziali dx_1 e dx_2 devono essere tali da mantenere la funzione inalterata al valore c . Riordinando la (2.8) si ottiene:

$$\frac{dx_2}{dx_1} = -f_1/f_2 \quad (f_2 \neq 0) \quad (2.9)$$

Il rapporto dei differenziali sul lato di sinistra può essere interpretato come la pendenza del contorno in un punto, in quanto soddisfano il vincolo che i loro valori siano tali da lasciare il valore della funzione inalterato. Quindi la (2.9) mostra che possiamo valutare la pendenza di un contorno in un punto come x' in figura 4(a) direttamente con i valori delle derivate parziali della funzione in quel punto. L'equazione (2.9) ci permette pure di determinare la pendenza del contorno dai segni delle derivate della funzione. Se presentano lo stesso segno, l'inclinazione del contorno deve essere negativa come nella figura 4(a), mentre sarà positiva come nella 4(b) se le derivate hanno segno opposto.

Nella teoria dell'ottimizzazione la continuità dei propri contorni è un'importante proprietà di cui una funzione può o meno godere. Una seconda proprietà molto importante è quella della *concavità* dei contorni. Al fine di esaminarla, ci limitiamo anzitutto a funzioni le cui derivate f_1 e f_2 siano positive; da ciò segue che, dati due vettori x' e x , con $x > x' \Rightarrow f(x) > f(x')$. La figura 5 illustra la concavità dei contorni per una simile funzione. La proprietà può essere descritta nel seguente modo: si scelgano due punti come x' e x'' nella figura, che giacciono sullo stesso contorno. Cioè: $f(x') = f(x'') = c$. Si scelga un punto qualsiasi sulla retta che congiunge x' e x'' , come \bar{x} nella figura. Allora il contorno è detto concavo se:

$$f(\bar{x}) \geq f(x') = f(x'') = c \quad (2.10)$$

Detto a parole, una combinazione convessa di due qualsiasi punti di un contorno produce un valore della funzione almeno altrettanto alto, e giace quindi sullo stesso contorno o su uno più alto. Una funzione i cui contorni soddisfino questa definizione è detta *quasi-concava*. Le funzioni con contorni del tipo (a) e (b) in figura 4 sono quasi-concave, mentre quelle del tipo (c) non lo sono. Per capire questo fatto si noti che in (a) e (b) parte di ogni contorno passa attraverso la zona grigia a sud-ovest di \bar{x} , implicando che alcuni punti sul contorno hanno valori di x_1 e x_2 inferiori di \bar{x} . Poiché riducendo x_1 e x_2 si riduce il va-

lore della
minore su
non acca
contiene p
la funzio
scussione
oltre al le

Una ult
gura. Nel
che li cong
cando che
maggiore
concava. C
se prendia
che li con

$$f(\hat{x}) = f$$

Quindi un
è strettam

Si noti i
della funzi
so destra,
soddisfare
un proble
valore dell
simizzare
più alto p



lore della funzione ($f_1, f_2 > 0$), ne segue che il valore della funzione deve essere minore sul contorno che non in x , e quindi la (2.10) è soddisfatta. In (c) invece non accade così: parte del contorno giace nell'area a nord-est di \bar{x} e quindi contiene punti per cui sia x_1 che x_2 sono maggiori che in \bar{x} . Perciò il valore della funzione è superiore lungo il contorno. Si noti il ruolo giocato in questa discussione dall'ipotesi che le derivate parziali siano positive. In un esercizio più oltre al lettore è chiesto di generalizzare tale risultato.

Una ulteriore distinzione può essere tracciata confrontando la (a) e la (b) in figura. Nel primo caso è chiaro che per *ogni* coppia di punti sul contorno la linea che li congiunge giacerà sempre completamente al di sopra del contorno, implicando che il valore della funzione in un punto come \bar{x} sarà sempre *strettamente* maggiore di quello sul contorno. Una simile funzione è detta *strettamente quasi-concava*. Questo non è d'altra parte il caso della funzione in (b). Per esempio, se prendiamo i punti x_0 e x^* sul contorno allora un punto come \hat{x} sulla linea che li congiunge giace anch'esso sul contorno, e quindi:

$$f(\hat{x}) = f(x^*) = f(x_0) = c \quad (2.11)$$

Quindi una funzione che possiede simili contorni, sebbene quasi-concava, non è *strettamente* quasi-concava⁽⁴⁾.

Si noti infine come variazioni del valore della costante c cambino il contorno della funzione. Date $f_1, f_2 > 0$, la crescita di c sposta i contorni in figura 5 verso destra, poiché per ogni dato x_2 , un valore maggiore di x_1 è necessario per soddisfare l'equazione che definisce il contorno (e viceversa). Ne segue che in un problema di ottimizzazione più alto è il contorno raggiunto, maggiore è il valore della funzione obiettivo, cosicché *possiamo considerare lo scopo di massimizzare la funzione obiettivo equivalente al tentativo di portarsi sul contorno più alto possibile*.

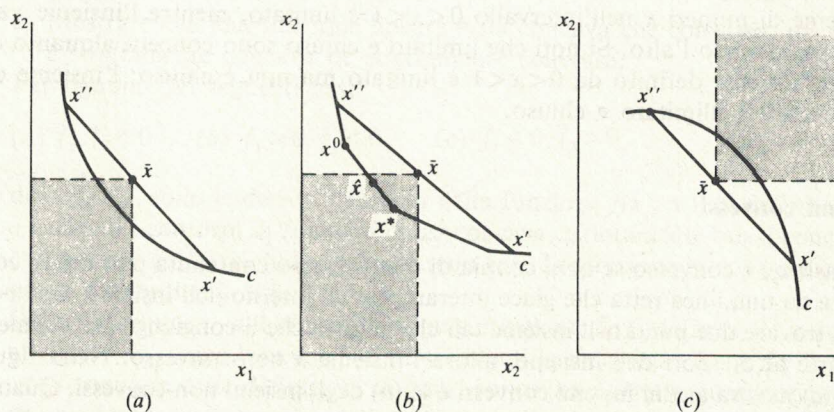


Fig. 5

Proprietà dell'insieme ammissibile

Vi sono quattro importanti proprietà degli insiemi che sono rilevanti dal punto di vista della teoria dell'ottimizzazione.

Insiemi non-vuoti

Un insieme è non-vuoto se contiene almeno un elemento, l'insieme vuoto essendo l'insieme senza elementi. Si ricordi che l'insieme ammissibile in un problema di ottimo è l'insieme dei punti o vettori x che soddisfano i vincoli. Un insieme ammissibile vuoto implica che non esiste nessuno di tali punti: i vincoli sono tali da escludere tutte le possibili soluzioni. Se i vincoli possono essere soddisfatti da almeno un punto, allora l'insieme ammissibile è non-vuoto.

Insiemi chiusi

Un insieme è chiuso se *tutti* i punti dei suoi confini sono elementi dell'insieme stesso. Quindi l'insieme di numeri x nell'intervallo $0 \leq x \leq 1$ è chiuso, mentre gli insiemi definiti per gli intervalli $0 < x < 1$ e $0 \leq x < 1$ non lo sono. Un esercizio qui sotto mostra la stretta relazione tra l'esistenza di disuguaglianze *deboli* nei vincoli di un problema e la chiusura dell'insieme ammissibile.

Insiemi limitati

Un insieme è limitato quando non è possibile procedere all'infinito pur rimanendo all'interno dell'insieme. In altre parole, sarà sempre possibile includere un insieme limitato in una sfera di dimensioni sufficientemente ampie. Perciò l'insieme di numeri x nell'intervallo $0 < x < 1$ è limitato, mentre l'insieme $x \geq 0$ è illimitato verso l'alto. Si noti che limitato e chiuso sono concetti alquanto distinti: l'insieme definito da $0 < x < 1$ è limitato ma non è chiuso; l'insieme dei valori $x \geq 0$ è illimitato e chiuso.

Insiemi convessi

Un insieme è convesso se *ogni* coppia di punti in esso contenuta può essere congiunta da una linea retta che giace interamente all'interno dell'insieme. Se è possibile trovare due punti nell'insieme tali che la linea che li congiunge giace almeno in parte al di fuori dell'insieme, allora l'insieme è non-convesso. Nella figura 6, la (a) mostra alcuni insiemi convessi e la (b) degli insiemi non-convessi. Quando due *qualsiasi* punti di confine di un insieme convesso siano uniti da una linea, se l'intera linea a eccezione dei suoi estremi giace entro l'insieme, allora que-

 x_2

st'ultimo è
la linea che
questo non
strettamen

Esercizi

1. a) Pro
 - b) Pro
il va
punti
 - c) Mos
è str
- 2.* Per og

(a) f_1, f_2

dove f_1
e tracci
e non-q

- nuti i c
- 3.* Dato
migliore
tete for

di una

- 4.* Servend