Nozioni Elementari di Matematica
Corso di Economia Politica (P/Z)

Prof. Giorgio Colacchio
Facoltà di Giurisprudenza
Università del Salento

(A.A. 2013-2014)

Avvertenza. A seguito di diverse richieste da parte degli studenti, rendo disponibili queste vecchie dispense che utilizzavo in alcuni corsi di Economia Politica svariati anni fa. Non sono ancora riuscito a riprodurre *ex novo* tutte le immagini (cosa che si rendeva necessaria, dal momento che il documento è stato appunto scritto molto tempo fa) ed alcune di queste, pertanto, presentano ancora una risoluzione grafica non ottimale: cercherò a tempo debito di risolvere questo problema.

È quasi inutile sottolineare che le nozioni – *molto elementari* – contenute in queste dispense dovrebbero *già* far parte del bagaglio di conoscenze di qualunque studente che abbia intrapreso un corso di studi universitari, a prescindere quindi dal percorso scolastico di provenienza. Pertanto queste dispense non sono assolutamente *obbligatorie* e non dovrebbero nemmeno rivelarsi *necessarie* ai fini del superamento dell'esame di Economia Politica. Possono però risultare utili per tutti coloro che, per una ragione o l'altra, abbiano bisogno di “rinverdire” alcune nozioni basilari di matematica necessarie per una adeguata comprensione del programma d'esame.

1. Alcune nozioni fondamentali sulla teoria degli insiemi

Nel trattare alcuni aspetti dell'algebra è utile servirsi della notazione e della terminologia della teoria degli insiemi. In matematica la parola 'insieme' è usata per indicare una collezione di oggetti considerata come un'entità singola. Si pensi ad esempio al concetto di "folla", "squadra", "elettorato": i singoli oggetti di ognuno di queste collettività sono chiamati elementi dell'insieme, e si dicono che appartengono a questo insieme o equivalentemente che sono contenuti in esso. L'insieme, a sua volta, contiene o è costituito dai suoi elementi. Un modo molto semplice ed intuitivo per fare riferimento ad un generico insieme è quello di pensare ad una scatola con dentro degli oggetti.
Gli insiemi di solito sono denotati con le lettere maiuscole: A, B, C, … e i corrispondenti elementi sono indicati con le lettere minuscole: a, b, c, …. Si usa la speciale notazione

$$a \in A$$

per indicare che "a è un elemento di A" o "a appartiene ad A". Se a non appartiene ad A scriveremo $a \notin A$. Un modo piuttosto consueto di scrivere un insieme è quello di riportare i suoi elementi tra parentesi graffe; per esempio l'insieme A delle facce pari di un dado può essere indicato con[1]

$$A \equiv \{2, 4, 6\}$$

Un insieme di particolare interesse è il cosiddetto insieme vuoto, ovvero un insieme che non contiene alcun elemento. L'insieme vuoto viene indicato con il simbolo $\varnothing$ e può essere intuitivamente interpretato come una scatola vuota. Il primo concetto fondamentale che mette in relazione un insieme con un altro è l'uguaglianza tra due insiemi:

Definizione di uguaglianza tra due insiemi. Due insiemi A e B sono detti uguali (o identici) quando consistono esattamente degli stessi elementi, nel qual caso scriviamo A=B. Se uno degli insiemi contiene un elemento non contenuto nell'altro, diciamo che gli insiemi non sono uguali e scriviamo

$$A \neq B.$$

[1] Conformemente alla letteratura corrente useremo un simbolo particolare ≡ (in luogo di =) per indicare uguaglianze che derivano, come in questo caso, da *definizioni*, ovvero quando il membro di «sinistra» prende significato da quello di «destra» (altri simboli che si usano in luogo di ≡ sono: $\stackrel{\text{def}}{=}$ e $\triangleq$). Il simbolo ≡ si legge, quindi, *è per definizione*.

Esempio 1. In conformità con questa definizione, i due insiemi $A \equiv \{2, 4, 6, 8\}$ e $B \equiv \{2, 8, 6, 4\}$ sono uguali, dal momento che entrambi consistono degli elementi 2, 4, 6, 8. L'esempio mostra che nella definizione di uguaglianza tra due insiemi è irrilevante l'ordine con cui gli elementi degli insiemi appaiono.

Esempio 2. Gli insiemi $A \equiv \{2, 4, 6, 8\}$ e $B \equiv \{2, 2, 4, 4, 6, 8\}$ sono uguali, anche se nell'insieme B ognuno dei due elementi 2 e 4 è elencato due volte. Entrambi gli insiemi contengono i quattro elementi 2, 4, 6, 8 e nessun altro; perciò, in base alla definizione precedente, dobbiamo dire che A e B sono uguali. Un esempio analogo è l'insieme delle lettere contenute nella parola Mississippi, che è uguale all'insieme $P \equiv \{M, i, s, p\}$.

Notare che l'insieme $\varnothing$ è diverso dall'insieme $\{\varnothing\}$, cioè $\varnothing \neq \{\varnothing\}$ (una scatola vuota è qualcosa di diverso rispetto ad una scatola che contiene una scatola vuota).

Da un insieme dato A possiamo ottenere nuovi insiemi, detti sottoinsiemi di A. In generale abbiamo la seguente definizione:

Definizione di sottoinsieme. Un insieme A è detto sottoinsieme di un insieme B, e si scrive $A \subseteq B$ quando ogni elemento di A appartiene anche a B. Si dice anche che A è contenuto in B (o che B contiene A). La relazione $\subseteq$ si chiama inclusione tra insiemi.

In base a tale definizione possiamo dedurre che l'insieme vuoto $\varnothing$ è un sottoinsieme di ogni insieme. Osservare inoltre che l'affermazione $A \subseteq B$ non esclude la possibilità che $B \subseteq A$. In effetti possiamo avere contemporaneamente $A \subseteq B$ e $B \subseteq A$, ma questo avviene solo se A e B hanno gli stessi elementi. In altre parole diremo che

$$A = B \text{ se e solo se } A \subseteq B \text{ e } B \subseteq A$$

Questo risultato è una conseguenza immediata delle precedenti definizioni di uguaglianza e inclusione. Se $A \subseteq B$ ma $A \neq B$, allora diremo che A è un sottoinsieme proprio di B; indichiamo questa relazione scrivendo $A \subset B$.

Dati due insiemi A e B possiamo ricavare un nuovo insieme, detto unione di A e B. Questo nuovo insieme è indicato con il simbolo

$$A \cup B$$

e si legge " A unione B".

Definizione di unione tra due insiemi. L'insieme unione $A \cup B$ è definito come l'insieme degli elementi contenuti in A, in B o in entrambi. In altre parole, $A \cup B$ è l'insieme di tutti gli elementi che appartengono ad almeno uno degli insiemi A e B.

Un altro insieme che possiamo ricavare a partire dai due insiemi A e B è l'intersezione tra A e B. In simboli

$$A \cap B$$

si legge "A intersezione B".

Definizione di intersezione tra due insiemi. L'insieme $A \cap B$ è definito come l'insieme degli elementi comuni sia ad A sia a B. Se i due insiemi A e B non anno elementi in comune la loro intersezione è l'insieme vuoto $\varnothing$ e in tal caso A e B sono detti disgiunti: $(A \cap B) = \varnothing$.

Definizione di differenza tra due insiemi. Se A e B sono due insiemi, la differenza $(A - B)$ (chiamata anche complemento di B relativo ad A) è definita come l'insieme di tutti gli elementi di A che non sono contenuti in B.

Le operazioni di unione e intersezione hanno molte analogie formali (ma anche molte differenze) con le operazioni di addizione e moltiplicazione dei numeri. Per esempio abbiamo le seguenti proprietà

Proprietà commutativa: $(A \cup B) = (B \cup A)$; $(A \cap B) = (B \cap A)$

Proprietà associativa: $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$

Proprietà distributiva: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Inoltre, è utile tenere presente che valgono le seguenti proprietà:

(P1)	$(A \cup A) = A$	;	(P2)	$(A \cap A) = A$
(P3)	$(A \cup \varnothing) = A$	;	(P4)	$(A \cap \varnothing) = \varnothing$
(P5)	$A \cup (A \cap B) = A$	;	(P6)	$A \cap (A \cup B) = A$
(P7)	$A - (B \cap C) = (A - B) \cup (A - C)$			

Definizione di prodotto cartesiano di due o più insiemi.

Dati due insiemi non vuoti A e B, l'insieme costituito dalle coppie ordinate (a, b), $a \in A$, $b \in B$, si chiama prodotto cartesiano (o semplicemente prodotto) dei due insiemi A e B e si indica con $A \times B$. Cioè:

$$A \times B \equiv \{(a, b);\ a \in A,\ b \in B\}.$$

Si badi bene che l'ordine in cui a, b ed A, B sono scritti in (a, b) ed $A \times B$ è essenziale in quanto la coppia (a, b) è diversa dalla coppia (b, a), e quindi se $B \neq A$ allora $A \times B \neq B \times A$. In altri termini una coppia è qualcosa di diverso da un insieme in quanto nella prima è significativo e distintivo l'ordine in cui gli elementi sono scritti, mentre nel secondo tale ordine, come sappiamo, non ha alcun rilievo, cioè: $(a, b) \neq (b, a)$ mentre $\{a, b\} = \{b, a\}$.
È evidente inoltre che nulla vieta di considerare il prodotto cartesiano di un insieme con se stesso, ovvero l'insieme: $A \times A \equiv \{(a, b);\ a \in A,\ b \in A\}$, costituito da tutte le possibili coppie ordinate che si possono determinare dagli elementi di A. In analogia con le proprietà dei numeri reali, generalmente si indica: $A \times A \equiv A^2$.

Esempio 3: dati i due insiemi: $A \equiv \{2, 4, 6\}$, $P \equiv \{M, i, s, p\}$, il prodotto cartesiano $A \times P$ sarà l'insie-

me:

$$A \times P = \{(2, M), (2, i), (2, s), (2,p), (4\ M), (4, i), (4, s), (4,p), (6, M), (6, i), (6, s), (6,p)\}.$$

Il prodotto cartesiano dell'insieme A con se stesso sarà invece l'insieme:

$$A \times A = \{(2, 2), (2, 4), (2, 6), (4,2), (4, 4), (4, 6), (6, 2), (6,4), (6, 6)\}.$$

È intuitivo estendere la definizione di **prodotto cartesiano a più di due insiemi**. Dati infatti tre insiemi non vuoti A , B e C, potremmo considerare il prodotto cartesiano tra l'insieme $A \times B$ e C, ottenendo l'insieme: $A \times B \times C$, e quindi considerare il prodotto cartesiano tra quest'ultimo insieme ed una altro insieme non vuoto D, ottenendo il nuovo insieme: $A \times B \times C \times D$, ecc. . Quindi, più in generale:

dati n insiemi non vuoti: $A_1, A_2, \ldots A_i \ldots A_n$, l'insieme costituito dalle n^{ple} *ordinate* $(a_1, a_2, \ldots a_i, \ldots a_n)$, $a_i \in A_i$, i = 1,2, ... n, si chiama prodotto cartesiano degli n insiemi e si indica con: $A_1 \times A_2 \ldots \times A_i \times A_n$. Evidentemente anche qui l'*ordine* in cui a_i e A_i sono scritti è essenziale.

Alcuni insiemi numerici

Sino a questo punto ci siamo occupati di insiemi "generici". Non abbiamo cioè specificato in modo chiaro la natura degli insiemi che stavamo trattando. Insiemi particolarmente importanti sono quelli contenenti numeri. In queste righe cercheremo di introdurre l'insieme dei numeri reali, partendo dall'insieme dei numeri interi.

L'insieme dei *numeri interi positivi*, detto anche *insieme dei numeri naturali*, è l'insieme di numeri più intuitivo. Per definire tale insieme partiamo dal numero 1. Il numero 1 + 1 è indicato con 2, il numero 2 + 1 è indicato con 3 e così via. I numeri 1, 2, 3, ..., ottenuti in questo modo, addizionando ripetutamente 1 al numero precedente, sono tutti positivi e sono detti interi positivi. Unendo per convenzione a tali numeri il numero 0, l'insieme dei numeri interi positivi può essere indicato con il simbolo:

$$\mathbf{N} \equiv \{0, 1, 2, \ldots\}$$

L'insieme dei numeri interi positivi ci permette quindi di "contare" oggetti o cose. Osservate che aggiungendo ogni volta il numero 1 ad un numero intero positivo precedente, si può anche contare

“all’infinito”. Gli opposti dei numeri interi positivi sono detti interi negativi e l’unione di **N** con gli interi negativi forma l’insieme:

$$\mathbf{Z} \equiv \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$$

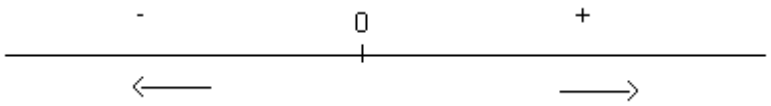
detto *insieme dei numeri interi*. Osserviamo che se $z \in \mathbf{Z}$, e $z \neq 0$, allora $(-z) + z = 0$, cioè la somma di un numero intero con il suo opposto è pari a zero. Nel caso dei numeri interi sarà possibile contare sia “in avanti”, che “all’indietro”. In altri termini, se ad esempio al numero intero -15437 aggiungiamo ripetutamente il numero intero (-1), otteniamo un numero sempre più piccolo in quanto negativo, ma più grande in valore assoluto[2].

A partire dall’insieme dei numeri **Z**, è possibile costruire l’insieme contenente tutti i rapporti (o frazioni) di interi del tipo: (a/b); tale insieme è detto *insieme dei numeri razionali* e viene indicato con

$$\mathbf{Q} \equiv \{\text{insieme di tutti i rapporti } a/b, \text{ dove } a \in \mathbf{Z} \text{ e } b \in \mathbf{Z} \text{ e } b \neq 0\}$$

E’ facile capire che sia **N** che **Z** sono sottoinsiemi di **Q**. Ad esempio, il numero (2/3) è un numero razionale in quanto è espresso come rapporto dei numeri interi 2 e 3; per indicare questo fatto scriveremo: $(2/3) \in \mathbf{Q}$. Anche il numero $(-15/7) \in \mathbf{Q}$, così come $7 = (7/1) \in \mathbf{Q}$.

Continuando la nostra discussione, possiamo osservare che l’insieme dei numeri razionali **Q** è un insieme “abbastanza ampio”, nel senso che contiene una gamma di numeri molto ampia. Ad ogni modo, per quanto uno tenti di sforzarsi, è impossibile esprimere il numero $\sqrt{2}$ come frazione di numeri interi. In altri termini, $\sqrt{2} \notin \mathbf{Q}$, e lo stesso discorso vale ad esempio per il famoso numero “pi greco”: $\pi = 3.14....$ I numeri che hanno questa caratteristica, cioè quella di non poter essere espressi come frazioni di numeri interi, sono detti *numeri irrazionali*. L’unione dei numeri irrazionali e dei numeri razionali genera un nuovo insieme di numeri molto interessante, cioè *l’insieme dei numeri reali*, che indichiamo con **R**. La proprietà più importante dei numeri reali **R** è che questi possono essere interpretati geometricamente come punti di una retta:



[2] Il *valore assoluto* di un numero *x*, che si indica con $|x|$, è dato da: $|x| = x$, se $x > 0$; $|x| = -x$ se $x < 0$.

nel senso che ad ogni punto della retta rimane associato un numero reale, e viceversa, ad ogni numero reale resta associato un punto della retta. È sufficiente solo indicare "l'origine" di tale retta attraverso il numero "zero" (0), dopodiché è facilissimo stabilire una corrispondenza tra punti della retta e numeri reali: per convenzione si suppone che i numeri "a destra" dello zero siano i numeri reali positivi, mentre quelli a sinistra dello zero i numeri reali negativi. Inoltre, poiché una retta è virtualmente infinita, dobbiamo pensare che fissato un qualsiasi numero reale, è sempre possibile trovare un altro numero reale più piccolo o più grande di quel prefissato numero.
Dalla discussione sviluppata sin ora possiamo facilmente comprendere che

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}$$

per cui se il numero 1 è un intero positivo, 1 sarà anche un numero intero, un numero razionale, ed un numero reale. Al contrario avremo che $\sqrt{2} \in \mathbf{R}$, ma $\sqrt{2} \notin \mathbf{Q}$, $\sqrt{2} \notin \mathbf{Z}$, $\sqrt{2} \notin \mathbf{N}$. In sostanza, l'insieme dei numeri reali costituisce l'insieme "più ampio" con il quale avremo a che fare.
Associate all'insieme **R** dei numeri reali, esistono due operazioni, chiamate addizione e moltiplicazione, tali che per ogni coppia x, y, di numeri reali possiamo ottenere la somma $(x + y)$ ed il prodotto $(x \cdot y)$. Valgono le seguenti proprietà:

Proprietà commutativa: $x + y = y + x$; $x \cdot y = y \cdot x$

Proprietà associativa: $x + (y + z) = (x + y) + z$; $x (y \cdot x) = (x \cdot y) \cdot z$

Proprietà distributiva: $x \cdot (y + z) = x \cdot y + x \cdot z$

Esistenza degli "elementi neutri": Esistono due numeri reali distinti, che indichiamo con 0 e con 1, tali che per ogni numero x reale abbiamo: $x + 0 = x$ e $x \cdot 1 = x$.

Esistenza dei numeri negativi: Per ogni numero reale x vi è un numero reale y tale che $x + y = 0$.

Esistenza dei reciproci: Per ogni numero reale $x \neq 0$, esiste un numero reale y tale che $x \cdot y = 1$.

Da queste caratteristiche dei numeri reali possiamo dedurre tutte le usuali leggi dell'algebra elementare. Particolarmente utili sono le seguenti semplici regole:

Regola 1: Dati a, b, c, numeri reali, se $a + b = a + c$, allora $b = c$.

Regola 1: Dati a, b, c, numeri reali, se $a \cdot b = a \cdot c$, allora $b = c$.

La *Regola 1* ci avverte che quando abbiamo un'espressione che coinvolge dei numeri reali ed in cui appare il segno di uguale: "=", tutto ciò che sommiamo (o sottraiamo) a sinistra del segno di uguale, deve essere a sua volta sommato (o sottratto) anche alla destra del segno di uguale; la *Regola 2* ci avverte che quando abbiamo un'espressione che coinvolge dei numeri reali ed in cui appare il segno di uguale: "=", tutto ciò che moltiplichiamo alla sinistra del segno di uguale deve anche essere moltiplicato alla destra del segno di uguale.

Focalizziamo per un momento l'attenzione sulla seguente espressione algebrica:

$$a \cdot x + b = 0$$

dove a, b, e x sono numeri reali. Supponiamo che i numeri a e b siano noti, e che il nostro obiettivo sia determinare il valore del numero reale x tale per cui l'espressione risulta verificata. Tale espressione è anche chiamata equazione lineare di primo grado. Come determiniamo il valore di x ? Se facciamo riferimento alla *Regola 1* e alle sua conseguenze, potremo senza dubbio scrivere:

$$a \cdot x + b + (-b) = 0 + (-b)$$

cioè abbiamo aggiunto il numero (-b) ad entrambi i termini a destra e a sinistra del segno di uguale. In questo modo otteniamo la nuova espressione algebrica:

$$a \cdot x = -b$$

che risulta perfettamente equivalente alla precedente. Ora, facendo riferimento alla *Regola 2* e alle sua conseguenze, possiamo moltiplicare il lato destro e sinistro di quest'ultima espressione, per una stessa quantità. Usando la quantità (1/a) otteniamo:

$$(1/a) \cdot a \cdot x = (1/a) \cdot (-b)$$

ovvero

$$x = -\frac{b}{a}$$

che è esattamente il valore di x che risolve l'equazione di primo grado scritta all'inizio.

Un breve cenno meritano quegli insiemi notevoli che si ottengono dal prodotto cartesiano ripetuto di **R** con se stesso, cioè: $\mathbf{R} \times \mathbf{R} \equiv \mathbf{R}^2$, $\mathbf{R} \times \mathbf{R} \times \mathbf{R} \equiv \mathbf{R}^2 \times \mathbf{R} \equiv \mathbf{R}^3$,..., $\mathbf{R}^{n-1} \times \mathbf{R} \equiv \mathbf{R}^n$. Come si può intuire, $\mathbf{R}^2$ è costituito da tutte le possibili coppie ordinate di numeri reali (a, b), $a \in \mathbf{R}$, $b \in \mathbf{R}$; $\mathbf{R}^3$ da tutte le possibili "triple" ordinate di numeri reali (a, b, c), $a \in \mathbf{R}$, $b \in \mathbf{R}$, $c \in \mathbf{R}$; quindi, in generale, $\mathbf{R}^n$ sarà costituito da tutte le possibili "ennuple" ordinate di numeri reali (a, b, c, ..., n), $a \in \mathbf{R}$, $b \in \mathbf{R}$, $c \in \mathbf{R}$ ... , $n \in \mathbf{R}$. L'insieme $\mathbf{R}^n$, $n \geq 2$, si chiama *spazio euclideo n - dimensionale*, mentre gli elementi di $\mathbf{R}^n$ si chiamano *punti* o *vettori* di $\mathbf{R}^n$.
Così come i numeri reali possono essere rappresentati come punti di una retta, le coppie ordinate di $\mathbf{R}^2$ possono essere rappresentate graficamente come punti di un piano ottenuto dall'intersezione della suddetta retta dei numeri reali con un'altra retta dei numeri reali alla prima perpendicolare: l'asse orizzontale (la prima retta) è detto *asse delle ascisse* e l'asse verticale (la seconda retta) è detto *asse delle ordinate*. Come è evidente, il piano viene a suddividersi in quattro quadranti: ruotando in senso antiorario, il primo è formato da tutte le coppie (a, b), $a \in \mathbf{R}$, $b \in \mathbf{R}$, con $a \geq 0$ e $b \geq 0$; il secondo da tutte le coppie (a, b), $a \in \mathbf{R}$, $b \in \mathbf{R}$, con $a \leq 0$ e $b \geq 0$; il terzo da tutte le coppie (a, b), $a \in \mathbf{R}$, $b \in \mathbf{R}$, con $a \leq 0$ e $b \leq 0$; il quarto da tutte le coppie (a, b), $a \in \mathbf{R}$, $b \in \mathbf{R}$, con $a \geq 0$ e $b \leq 0$; il punto in cui le due rette si intersecano è individuato dalla coppia: 0 = (0, 0). Data la coppia (a, b) l'elemento a (il primo elemento della coppia) viene detto *ascissa* e l'elemento b (il secondo elemento della coppia) viene detto *ordinata*; l'ascissa e l'ordinata rappresentano poi le *coordinate* del punto $P \equiv (a, b)$, ovvero del punto individuato dalla coppia (a, b): generalmente si dice che l'ascissa è la *prima coordinata* e l'ordinata la *seconda coordinata*. Nel grafico sotto sono rappresentati, a titolo esemplificativo, quattro punti di $\mathbf{R}^2$: il punto A ha prima coordinata positiva e seconda nulla, il punto B ha prima coordinata negativa e seconda positiva, il punto C ha entrambe le coordinate negative ed il punto D ha prima coordinata positiva e seconda coordinata negativa.

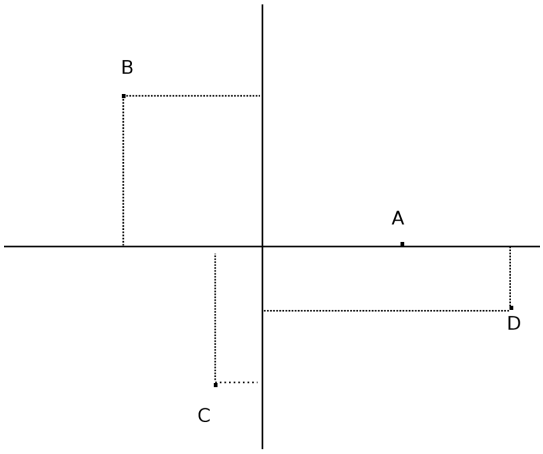


Figura 1

Anche per gli elementi di $\mathbf{R}^3$ si può tentare di fornire una rappresentazione geometrica intuitiva: le triple ordinate di $\mathbf{R}^3$ possono essere infatti rappresentate graficamente come punti di uno *spazio tri-dimensionale* ottenuto tracciando una retta, perpendicolare al piano bidimensionale precedentemente visto, passante per il punto 0. Nella seguente figura è rappresentato il punto A ≡ (a, b, c) con entrambe le coordinate positive (si noti che nella figura i numeri reali negativi giacciono sulla parte tratteggiata delle rette; inoltre anche qui l'intersezione delle tre rette individua il punto 0 di coordinate 0 ≡ (0, 0, 0)):

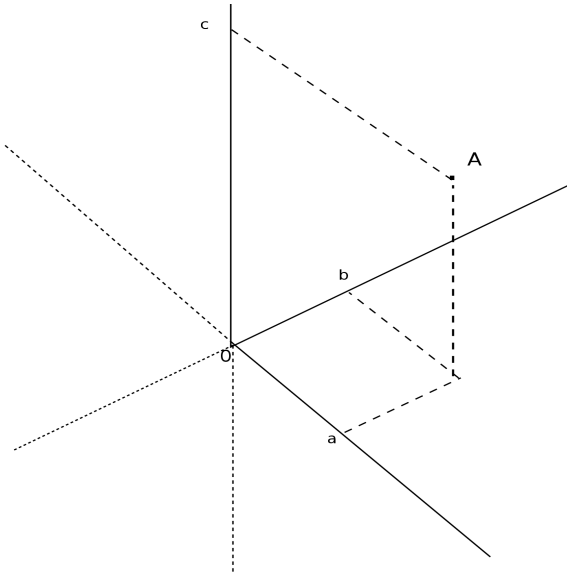


Figura 2

2. Relazioni e funzioni

Dati due insiemi A e B, si definisce funzione (o *trasformazione* o *mappa*) *f* da A a B, e si indica con $f : A \rightarrow B$ (si legge «effe è una funzione da A a B») una relazione[3], se esiste, che associa *ad ogni* elemento $x \in A$ *uno ed un solo* elemento $y \in B$.

Proprio perché $x \in A$ può essere scelto arbitrariamente mentre il corrispondente $y \in B$ risulta determinato univocamente attraverso *f*, si suole definire usualmente x *variabile indipendente* ed y *variabile dipendente* (o *immagine* di x attraverso *f*), e si scrive[4]:

$$y = f(x) \text{ (si legge: «y è uguale ad effe di x»)}$$

L'insieme A si chiama *dominio* o *insieme di definizione* della funzione *f*, l'insieme B si chiama *codominio* della funzione, mentre l'insieme formato da tutti quegli elementi di B che risultano dalla trasformazione $y = f(x)$ *per ogni* $x \in A$, si chiama *immagine* di A e viene solitamente indicato con $f(A)$.

L'immagine $f(A)$ è ovviamente un sottoinsieme di B, ovvero $f(A) \subseteq B$, è può coincidere o meno con tutto B. Nel caso in cui $f(A) = B$ si dice che *f* è una funzione *suriettiva*, ovvero una funzione da A *su* B.

Un importante concetto è quello di funzione *iniettiva*.

Una funzione $f : A \rightarrow B$ si dice iniettiva se $x_1 \neq x_2$ implica che $f(x_1) \neq f(x_2)$, *per ogni* $x_1, x_2 \in A$, $x_1 \neq x_2$, ovvero se *f* trasforma elementi distinti di A in elementi distinti di B.

Intuitivamente, quindi, una funzione si dirà iniettiva quando l'immagine di *ogni* elemento del dominio sarà diversa dall'immagine di *qualunque altro* elemento del dominio.

I seguenti esempi dovrebbero chiarire quanto fin qui detto.

Esempio 1.

Sia l'insieme A costituito dai numeri naturali compresi tra 1 e 6, e l'insieme B costituito dalle parole: *neo*, *rupe*, *seta*, *testa*, *sasso*, *sapore*, cioè:

$$A \equiv \{1, 2, 3, 4, 5, 6\}; \quad B \equiv \{\textit{neo}, \textit{rupe}, \textit{seta}, \textit{testa}, \textit{sasso}, \textit{sapore}\}$$

---

[3] In termini intuitivi una *relazione R* da un insieme A ad un insieme B è una «regola» attraverso la quale elementi di A vengono associati ad elementi di B. Il concetto di relazione è quindi molto più *generale* di quella di funzione, in quanto non esclude che alcuni elementi di A (al limite tutti) possano essere associati a *più* elementi di B, e che inoltre alcuni elementi di A (al limite tutti) possano non essere associati ad alcun elemento di B. Un esempio chiarificatore verrà dato in seguito.

[4] Per denotare le funzioni si usano generalmente le lettere minuscole *f*, *g*, *h*, ecc.. Ovviamente nel caso in cui si abbia $y = h(x)$, si leggerà «y è uguale ad acca di x» ecc..

e definiamo la relazione *R* che associa ad ogni elemento a $\in$ A quegli elementi di B il cui numero di lettere è pari ad a: in altre parole diremo che a $\in$ A è in relazione con b $\in$ B se b è una parola composta da una numero di lettere pari ad a: in tal caso diremo che la coppia (a; b) soddisfa la suddetta relazione.

Questa relazione, come si può comprendere, non è una funzione, in quanto *non tutti* gli elementi di A sono in relazione con qualche elemento di B (i numeri 1 e 2 non sono infatti in relazione con alcun elemento di B) ed inoltre alcuni elementi di A (i numeri 4 e 5) sono in relazione con *più di un* elemento di B (le coppie (4; *seta*), (4; *rupe*) così come le coppie (5; *testa*), (5; *sasso*) soddisfano *R* e perciò sia il numero 4 che il numero 5 sono in relazione con *due* elementi di B).

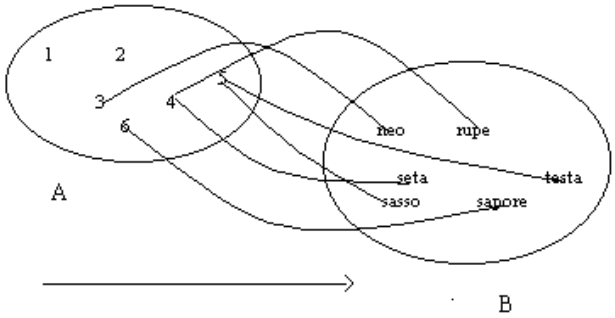


Figura 3

Esempio 2.

La relazione che associa ad ogni numero intero il suo quadrato è evidentemente una funzione, definita dall'insieme dei numeri interi all'insieme dei numeri naturali (essendo il quadrato di un numero intero sempre positivo), cioè:

$$f\colon \mathbf{Z} \to \mathbf{N}, \text{ definita da: } y = x^2,\ x \in \mathbf{Z},\ y \in \mathbf{N}.$$

Come si può notare tale funzione associa ad *ogni* elemento di **Z** *uno ed un solo* elemento di **N**. È inoltre evidente che *f* non è suriettiva (in quanto $f(\mathbf{Z}) \neq \mathbf{N}$, ovvero $f(\mathbf{Z}) \subset \mathbf{N}$) e non è iniettiva (numeri opposti hanno infatti lo stesso quadrato e quindi $f(1) = f(-1) = 1$, $f(2) = f(-2) = 4$ ecc.).

Esempio 3.

La funzione $f\colon \mathbf{Z} \to \mathbf{N}$ che associa ad ogni numero intero il suo valore assoluto, cioè:

$$y = |x|, \quad x \in \mathbf{Z}, y \in \mathbf{N}$$

è suriettiva, seppure non è iniettiva (perché?).

Esempio 4.

La funzione f: $\mathbf{N} \rightarrow \mathbf{P}$, ove $\mathbf{P}$ è l'insieme dei numeri pari positivi: $\mathbf{P} \equiv \{0, 2, 4, 6, 8...\}$, definita da:

$$y = 2x, \quad x \in \mathbf{N}, y \in \mathbf{P}$$

è suriettiva ed iniettiva (mentre la funzione f: $\mathbf{Z} \rightarrow \mathbf{P}$, definita da $y = 2|x|$ sarà suriettiva ma non iniettiva...perché?).

Possiamo adesso porci un interessante problema. Come abbiamo visto, attraverso una funzione ogni elemento di un insieme «di partenza» (il dominio) viene trasformato in uno ed un solo elemento di un insieme di «arrivo» (il codominio). È intuitivo, d'altra parte, che una volta nota la funzione sarà sempre possibile percorrere il cammino a ritroso: partire cioè da un elemento del codominio risultato dalla «trasformazione» tramite f, per risalire all'elemento o agli elementi (del dominio) di partenza. Data la funzione f: $A \rightarrow B$, *se* ad ogni elemento dell'immagine $f(A)$ corrisponde *uno ed un solo* elemento del dominio A, potremo esprimere questa relazione $f(A) \rightarrow A$ attraverso una nuova funzione, chiamata *funzione inversa* di f ed indicata solitamente con f^{-1}, ovvero: f^{-1}: $f(A) \rightarrow A$ (in tal caso si dice che la funzione f è *invertibile*).

Da quanto detto è evidente che una condizione *necessaria* per l'esistenza dell'inversa f^{-1} è che f sia una funzione iniettiva: soltanto in questo caso, infatti, ad ogni elemento dell'immagine corrisponderà *uno ed un solo* elemento del dominio. Se si ritorna all'*Esempio 2* si può verificare, infatti, come ogni $y \in f(\mathbf{Z})$, $y \neq 0$, sia "generato" da *due* elementi di $\mathbf{Z}$ (ad es. $4 \in f(\mathbf{Z})$ è generato da $x = 2^2$ e $x = -2^2$, $9 \in f(\mathbf{Z})$ da $x = 3^2$ e $x = -3^2$ ecc.), e le stesse considerazioni valgono per l'*Esempio 3*. Soltanto nell'ultimo caso (*Esempio 4*) ogni $y \in f(\mathbf{N}) = \mathbf{P}$ è univocamente riconducibile ad un *unico* elemento $x \in \mathbf{N}$ (ad es. $2 \in f(\mathbf{N})$ è generato da $x = 1$, $4 \in f(\mathbf{N})$ da $x = 2$ ecc.), e pertanto sarà possibile scrivere la funzione inversa[5]: f^{-1}: $\mathbf{P} \rightarrow \mathbf{N}$, definita da: $x = (\frac{1}{2})\, y$, $y \in \mathbf{P}$, $x \in \mathbf{N}$.

Per definire una funzione $f(A) \rightarrow B$ nessuna particolare restrizione, come abbiamo visto, è stata po-

[5] In questo caso, essendo f suriettiva, $f(\mathbf{Z}) = \mathbf{P}$.

sta sul dominio e sul codominio: sia A che B possono quindi essere degli insiemi molto generali i cui elementi non necessariamente dovranno essere così "semplici" come quelli fin qui considerati. La definizione di *prodotto cartesiano* di due o più insiemi ci consente di approfondire ed estendere la nozione di funzione.

Si consideri, per cominciare, il prodotto cartesiano dell'insieme U ≡ {1} (U quindi è l'insieme costituito dall'unico elemento 1) con l'insieme **N** dei numeri naturali precedentemente definito, cioè il nuovo insieme: A ≡ U × **N** = {(1,b); b ∈ **N** } costituito da tutte le possibili coppie (1, b) *per ogni* b ∈ **N**, e definiamo la relazione che ad ogni elemento (1, b) di A associa la somma (1+b). In tal modo, evidentemente, abbiamo "costruito" una funzione *f* : A → **N**, dal momento che la nostra relazione soddisfa la definizione (di funzione) precedentemente data: come si vede, infatti, attraverso *f* ogni elemento di A è associato ("trasformato" in) *uno ed un solo* elemento di **N**[6].

Esempio 5

Dati i due insiemi: U ≡ {1}, **N*** ≡ {1, 2, 3, 4,...} (**N*** è l'insieme dei numeri naturali maggiori di zero), la relazione che associa ad ogni elemento dell'insieme A ≡ U × **N*** = {(1,b); b ∈ **N*** }, l'elemento 1/b, è una funzione la cui immagine è un sottoinsieme di **Z**, cioè: *f* : A → **Z,** *f*(A) = 1/b, per ogni b ∈ **N***.

L'uso delle funzioni nell'ambito delle scienze economiche è generalmente motivato dalla necessità di esprimere il legame (ovvero la "dipendenza") che esiste tra due o più variabili, cercando quindi di esplicitare e rappresentare, in modo rigoroso, i comportamenti degli agenti economici e le relazioni causali che si instaurano tra di essi (nel tentativo più generale di definire in modo sintetico un "modello" che descriva al meglio gli aspetti economici oggetto di indagine): la costruzione di una tale "rappresentazione", se fedele, può così costituire un utile strumento per effettuare predizioni e più in generale per indirizzare l'azione dei *policy makers*. È evidente, comunque, che dietro rappresentazioni anche altamente "matematizzate" si nascondono sempre degli approcci teorici, ovvero delle visioni della realtà che, in particolare nel campo dell'economia, risultano spesso essere alternative (se non a volte fortemente confliggenti e quasi contrapposte): ciò per sottolineare che sottostanti alle funzioni, e più in generale ai modelli che incontreremo, giacciono sempre delle "teorie", e la valutazione del loro grado di adattamento è un compito ulteriore, spesso demandato alla verifica empirica condotta dagli statistici e più in particolare dagli econometrici.

[6] E possiamo quindi estendere a questo caso tutte le nostre precedenti considerazioni. In particolare si noti che *f* è iniettiva ma non suriettiva (perché?).

Esempio 6

Una delle più semplici funzioni che si incontrano in un corso di economia è quella che mette in relazione il livello dei risparmi (S) con il reddito (Y): S = f (Y), S, Y, $\in$ **R*** (**R*** è l'insieme dei numeri reali positivi), assumendo, conformemente alla *teoria keynesiana*, che all'aumentare di Y aumenti anche (in maniera più che proporzionale) S: soggetti che hanno un reddito più elevato tenderanno infatti a risparmiare di più di soggetti il cui reddito è basso, e ciò non solo in termini assoluti, ma anche in proporzione al loro reddito.

Secondo un'altra teoria, comunque, sostenuta in particolare dagli economisti *neoclassici*, più che il reddito la principale variabile da cui dipenderebbero i risparmi sarebbe il tasso di interesse i, ovvero: S = g (i), S, i, $\in$ **R***, in quanto più elevato è i, più alta è la remunerazione del risparmio e quindi maggiore sarà lo stimolo a risparmiare.

L'esempio appena presentato è utile in quanto ci ha mostrato *l'uso* delle funzioni per descrivere e spiegare i comportamenti economici, evidenziando inoltre come tale descrizione possa non essere univoca. È chiaro, evidentemente, che la scelta della prima o della seconda funzione implica un diverso approccio teorico, ovvero una diversa visione della realtà economica.

Per quanto le funzioni che fanno parte della "cassetta degli attrezzi" di un economista siano le più varie (e complesse) possibili, in un corso base di economia esse si restringono notevolmente, essendo sufficiente quasi sempre "adoperare" *funzioni reali di una variabile reale*, e cioè funzioni del tipo: $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$. Dovrebbe anzi essere intuitivo (si consideri l'esempio appena visto) che sull'insieme di definizione delle funzioni, così come sul loro codominio, opererà quasi sempre l'ulteriore restrizione di essere limitati ad **R*** (il sottoinsieme dei numeri reali positivi), non essendo significative, in genere[7], variabili economiche negative, ovvero: $f : A \rightarrow \mathbf{R}^*$, $A \subseteq \mathbf{R}^*$.

Tra queste funzioni elementari, quella la cui conoscenza è essenziale, essendo la più adoperata in un corso base di Economia, è la funzione *affine* (spesso chiamata anche funzione *lineare*) definita da:

$$y = mx + q, \qquad m, x, q \in \mathbf{R}$$

[7] Si tenga comunque conto che in analisi avanzate ciò non è più vero, considerandosi anche variabili economiche negative (un esempio intuitivo è quello dei beni "nocivi" (*noxious commodities*), si pensi alle scorie, per i quali si assumono prezzi negativi).

ove y è la variabile dipendente ed x la variabile indipendente.
Il termine m viene spesso definito *coefficiente angolare*, mentre q definisce *l'ordinata all'origine* della funzione. L'interpretazione geometrica di una funzione affine è particolarmente intuitiva, essendo rappresentata da una *retta* nel piano cartesiano [x, y]: una volta noti m e q (nell'ipotesi $m \neq 0$) è sufficiente assegnare arbitrariamente due diversi valori alla x (ad es. x_1, x_2) ottenendo i corrispondenti (diversi[8]) valori di y (ad es. y_1, y_2) e quindi tracciare la retta[9] passante per la coppia di punti (x_1, y_1), (x_2, y_2) così individuati nel piano cartesiano.
Ponendo $x = 0$ si ha $y = q$, che rappresenta il punto in cui la retta interseca l'asse delle ordinate[10] (da qui il nome con cui viene indicato q), mentre ponendo $y = 0$, nell'ipotesi sempre che m sia diverso da zero, si ha $x = - q/m$, che è il punto in cui la retta interseca l'asse delle ascisse[11]: questi due punti vengono spesso definiti le *intercette* (con gli assi cartesiani) della retta. Come si realizza immediatamente per $q \neq 0$: se $m > 0$ le intercette avranno sempre segno opposto (cioè una sarà negativa ed una positiva), mentre se $m < 0$ esse avranno sempre lo stesso segno (entrambe positive oppure entrambe negative). Per $q = 0$ ed $m \neq 0$ la retta invece passerà per l'origine: in tal caso le intercette saranno coincidenti, e pari entrambe a zero[12].
Per ultimo, se $m = 0$ e $q \neq 0$ (se anche q fosse nullo non potremmo evidentemente definire la funzione) l'equazione (1) si trasforma nella *funzione costante*: $y = q$.

La rappresentazione del legame esistente tra due variabili economiche, x ed y, attraverso una funzione: $y = f(x)$, $x, y \in \mathbf{R}$, ci permette di comprendere immediatamente la relazione quanti-qualitativa che sussiste tra le due variabili. Specificamente, la forma funzionale rifletterà il *tipo* di relazione (decrescente o crescente) che sussiste tra la x e la y, informandoci inoltre sull'*entità* della variazione della y conseguente ad una determinata variazione della variabile indipendente. Il primo aspetto concerne evidentemente gli aspetti *qualitativi* del legame funzionale: all'aumentare (ovvero al diminuire) di x, la y aumenta o diminuisce? Il secondo aspetto ci dà invece delle informazioni ulteriori di natura *quantitativa*, indagando sulla magnitudine della suddetta variazione: per una *determinata* variazione della x di *quanto* aumenta (ovvero diminuisce) la y? E supponendo di far variare la x sempre della stessa quantità, l'entità della variazione della y sarà costante o variabile lungo tutto il dominio della funzione?
In termini più rigorosi, adoperando il simbolo Δ per indicare, come di consueto, la variazione di una

[8] *Diversi* in quanto, come si può vedere, se $m \neq 0$ la funzione è iniettiva.
[9] Si può dimostrare infatti che per due punti distinti passa una ed una sola retta.
[10] Tale punto è individuato infatti dalle coordinate (0, q).
[11] Tale punto è individuato infatti dalle coordinate ($-m/q$, 0).
[12] Ovvero il punto (0, 0), che individua l'origine, soddisferà l'equazione della retta.

variabile[13], cioè: $\Delta x \equiv x_2 - x_1$, e definito *rapporto incrementale* il rapporto $\frac{\Delta y}{\Delta x}$, possiamo così riassumere la due questioni:

- Data la funzione: y = $f(x)$, $x, y \in \mathbf{R}$, se $\Delta x > 0$ (ovvero $\Delta x < 0$), quale sarà il segno di $\Delta y = f(x_2) - f(x_1)$ (cioè $\Delta y \gtreqless 0$) ?

- Data la funzione: y = $f(x)$, $x, y \in \mathbf{R}$, supposto un $\Delta x \equiv x_2 - x_1 = \theta =$ costante, quale sarà l'entità del rapporto incrementale $\frac{\Delta y}{\Delta x}$? E, più in generale, per *ogni* possibile coppia di valori: x_n , x_m, tale che $\Delta x = x_n - x_m = \theta$, quale sarà l'entità del corrispondente $\frac{\Delta y}{\Delta x}$?

Nel caso della funzione affine che abbiamo visto: $y = mx + q$, ($m, x, q \in \mathbf{R}$), la risposta a queste due domande è immediata, e dipende crucialmente dal coefficiente angolare.
Il *segno* di m stabilisce il tipo di legame che sussiste fra le due variabili x ed y: se $m > 0$ ($m < 0$) all'aumentare del valore della variabile indipendente aumenta (diminuisce) il valore della variabile dipendente. Il *valore* di m definisce invece l'entità della variazione Δy a fronte di un determinato Δx, essendo $\Delta y = m\Delta x$.

Data infatti la funzione affine: $y = mx + q$ ($m, x, q \in \mathbf{R}$) e considerati due punti del dominio: x_1, x_2, con $x_1 < x_2$, si ha:

1) $y_1 = mx_1 + q$; $y_2 = mx_2 + q$, da cui:

2) $y_2 - y_1 = mx_2 + q - (mx_1 + q) = m(x_2 - x_1)$. Essendo per ipotesi $x_2 - x_1 > 0$, si evince immediatamente che:

3) se $m > 0$ ($m < 0$) $\rightarrow$ $y_2 - y_1 > 0$ ($y_2 - y_1 < 0$).

Inoltre, definito al solito $\Delta x = x_2 - x_1$, dalla (2) si ricava immediatamente che:

4) $y_2 - y_1 = \Delta y = m(x_2 - x_1) = m\,\Delta x$, ovvero: $\Delta y = m\,\Delta x$, *qualunque* sia la variazione Δx ipotizzata, e

[13] Variazione che ovviamente potrà essere sia positiva (se $x_2 > x_1$) che *negativa* (se $x_2 < x_1$).

quindi (purché $\Delta x \neq 0$) $m = \dfrac{\Delta y}{\Delta x}$.

Da quanto visto dovrebbe ormai essere chiaro che *m* è un indicatore della *pendenza* della retta che rappresenta la funzione affine $y = mx + q$ nel piano cartesiano [*x*- *y*]. Misurando positivamente, per convenzione, in senso *antiorario* gli angoli, se $m > 0$ tale retta formerà un angolo compreso tra 0° e 90° con l'asse delle ascisse[14] e diremo che essa è *positivamente* inclinata, mentre per $m < 0$ essa formerà un angolo compreso tra 0° e -90° con l'asse delle ascisse[15] e diremo che è *negativamente* inclinata. È evidente, quindi, che quanto più il valore assoluto di *m* sarà elevato tanto più la retta sarà (positivamente o negativamente) inclinata e per $m \to +\infty$ ($m \to -\infty$) essa tenderà a diventare verticale, ovvero l'angolo formato con l'asse delle ascisse tenderà a +90° (-90°)[16]. Consegue quindi immediatamente che funzioni affini che presentano lo stesso coefficiente angolare saranno *parallele* (e coincideranno se anche il valore di *q* sarà lo stesso).

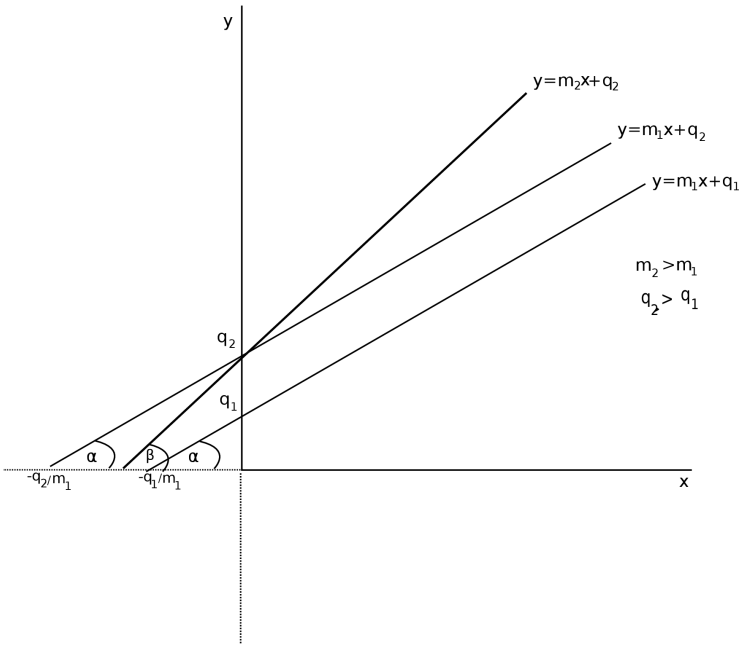


Figura 4

[14] Ovvero un angolo negativo compreso, in valore assoluto, tra 90° e 180°.
[15] Ovvero un angolo positivo compreso tra 90° e 180°.
[16] In termini rigorosi, *m* definisce la *tangente trigonometrica* dell'angolo che la retta forma con l'asse delle ascisse.

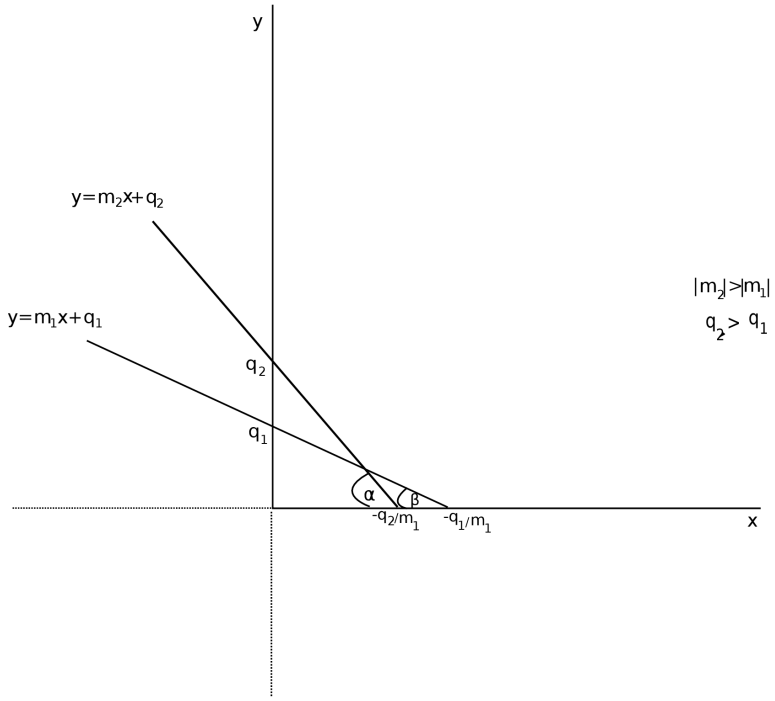


Figura 5

Consideriamo adesso una funzione generica $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$: $y = f(x)$, e tentiamo di rispondere alle precedenti due domande avvalendoci, sostanzialmente, dell'ispezione grafica. È evidente che, a meno di non considerare la funzione affine che abbiamo precedentemente esaminato, il rapporto incrementale non sarà in generale costante potendo inoltre il suo stesso segno variare lungo il dominio della funzione: in tal modo la *f* potrà incorporare ipotesi "più complicate" sul tipo di legame esistente tra la variabile indipendente e quella dipendente, esprimendo possibili modificazioni di natura sia quantitativa che qualitativa al variare della x.

Forniamo innanzitutto alcune semplici ed intuitive definizioni che aiutano notevolmente a semplificare la trattazione.

Data la funzione: $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$ definita da: $y = f(x)$, se *per ogni* coppia $x_1, x_2 \in A$, con $x_1 < x_2$, $f(x_2) \geq f(x_1)$, si dice che *f* è *crescente* su A; se al contrario *per ogni* coppia $x_1, x_2 \in A$, con $x_1 < x_2$, $f(x_2) \leq f(x_1)$, si dice che *f* è *decrescente* su A[17].

[17] Se valgono le disuguaglianze forti, ovvero se *per ogni* coppia $x_1, x_2 \in A$, con $x_1 < x_2$, $f(x_2) > f(x_1)$ ($f(x_2) < f(x_1)$) si dice che *f* è *strettamente crescente* (*strettamente decrescente*). È immediato verificare che la funzione affine precedentemente considerata: $y = mx + q$, sarà strettamente crescente se $m > 0$, strettamente decrescente se $m < 0$.

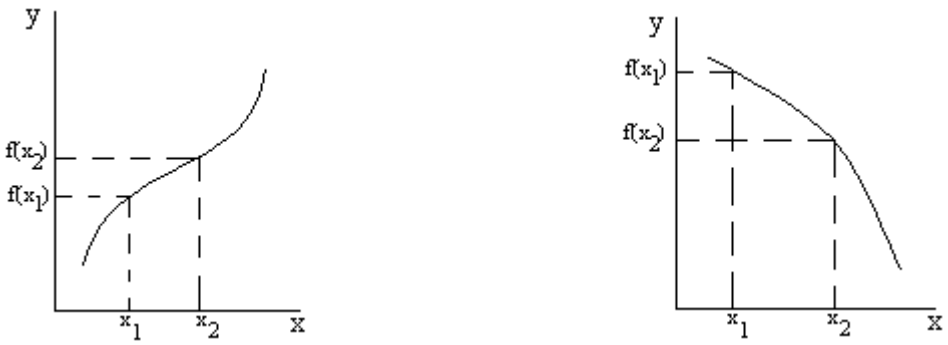


Figura 6

Nulla vieta, evidentemente, di restringere queste definizioni a sottoinsiemi di A, ovvero ad intervalli opportunamente prescelti del dominio. Data cioè la funzione $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$ definita da: $y = f(x)$, se *per ogni* coppia $x_1, x_2 \in B$ A, con $x_1 < x_2$, $f(x_2) \geq f(x_1)$, si dice che f è *crescente* su B; se al contrario *per ogni* coppia $x_1, x_2 \in B \subset A$, con $x_1 < x_2$, $f(x_2) \leq f(x_1)$, si dice che f è *decrescente* su B[18].

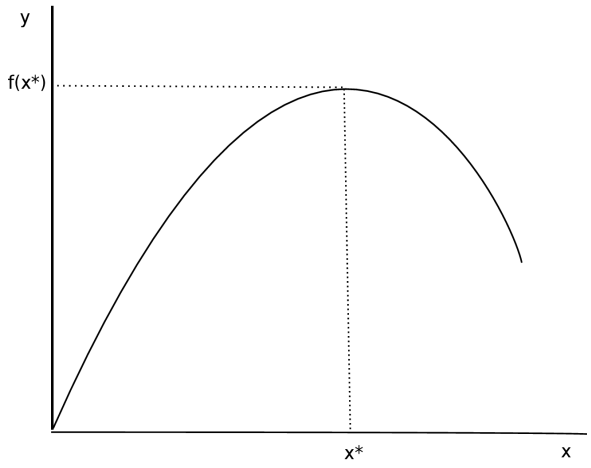


Figura 7

Consideriamo adesso la funzione strettamente crescente (in cui cioè per ipotesi $\frac{\Delta y}{\Delta x} > 0$)

$f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$, $y = f(x)$, rappresentata nel grafico seguente :

[18] Vale quanto affermato nella nota precedente nel caso dovessero valere le disuguaglianze forti. Come si può vedere nella figura 7, la f è strettamente crescente per $0 < x \leq x^*$ mentre è strettamente decrescente per $x > x^*$.

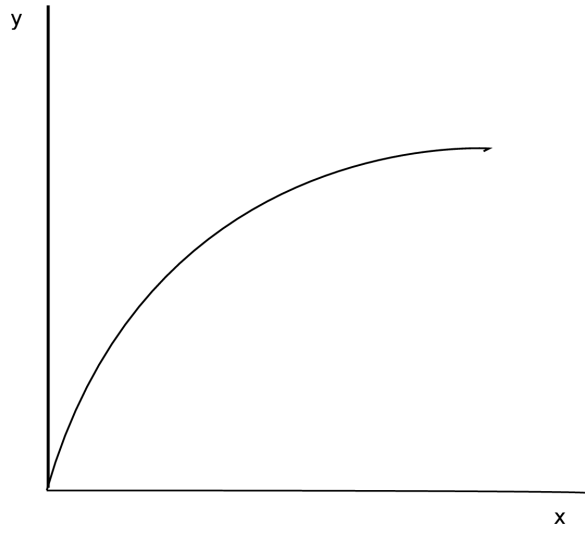


Figura 8

e studiamo, attraverso una semplice ispezione grafica, come varia il rapporto incrementale al variare della variabile indipendente. Il modo di procedere è, come al solito, quello di considerare un *determinato* incremento della variabile indipendente $\Delta x = x_n - x_m = \theta, x_n, x_m \in \boldsymbol{A}$, applicandolo successivamente lungo il dominio della funzione, al fine di determinare il corrispondente Δy.

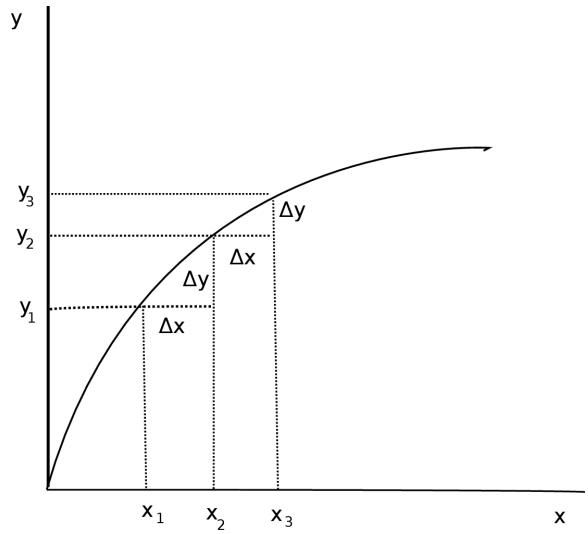


Figura 9

Come si può notare, il rapporto incrementale $\frac{\Delta y}{\Delta x}$, nel caso rappresentato in Fig. 9, è decrescente all'aumentare di *x* (per identici incrementi successivi Δx, il corrispondente Δy diminuisce progressivamente), indicando la curvatura della funzione. Da quanto precedentemente detto sulle rette, è

inoltre evidente come un *determinato* $\frac{\Delta y}{\Delta x}$ (con $\Delta x = x_n - x_m = \theta, e \Delta y = f(x_n) - f(x_m)$) altro non rappresenti che la pendenza (o meglio il coefficiente angolare) della retta passante per i punti di coordinate (x_n, $f(x_n)$), (x_m , $f(x_m)$), misurando quindi l'angolo che quest'ultima forma con l'asse delle ascisse. Come si può notare, per valori sempre più elevati della *x* tale angolo diminuisce continuamente:

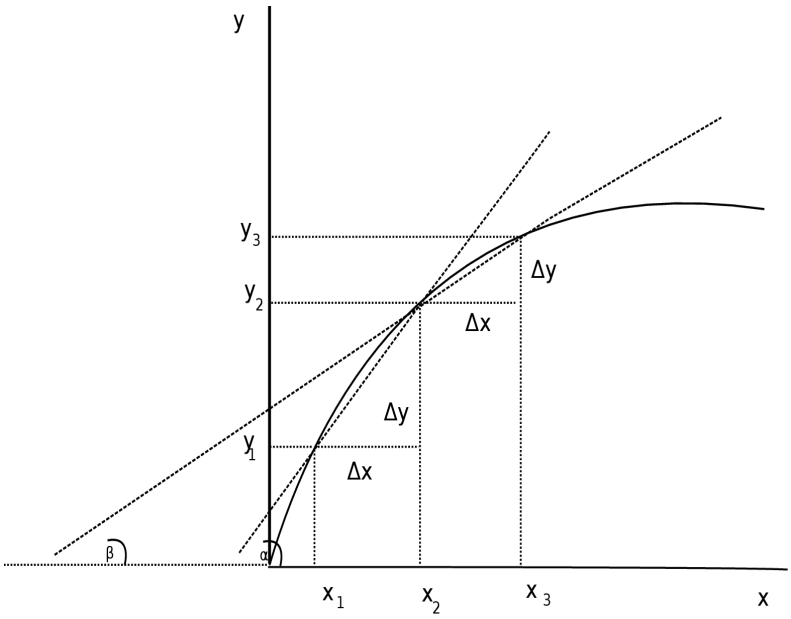


Figura 10

Nel caso invece della funzione (strettamente) crescente rappresentata nella figura seguente, il rapporto incrementale, come si può agevolmente notare, aumenta continuamente per valori sempre maggiori della variabile indipendente; detto in altri termini, gli angoli formati dalle rette con l'asse delle ascisse saranno sempre più ampi al crescere di x:

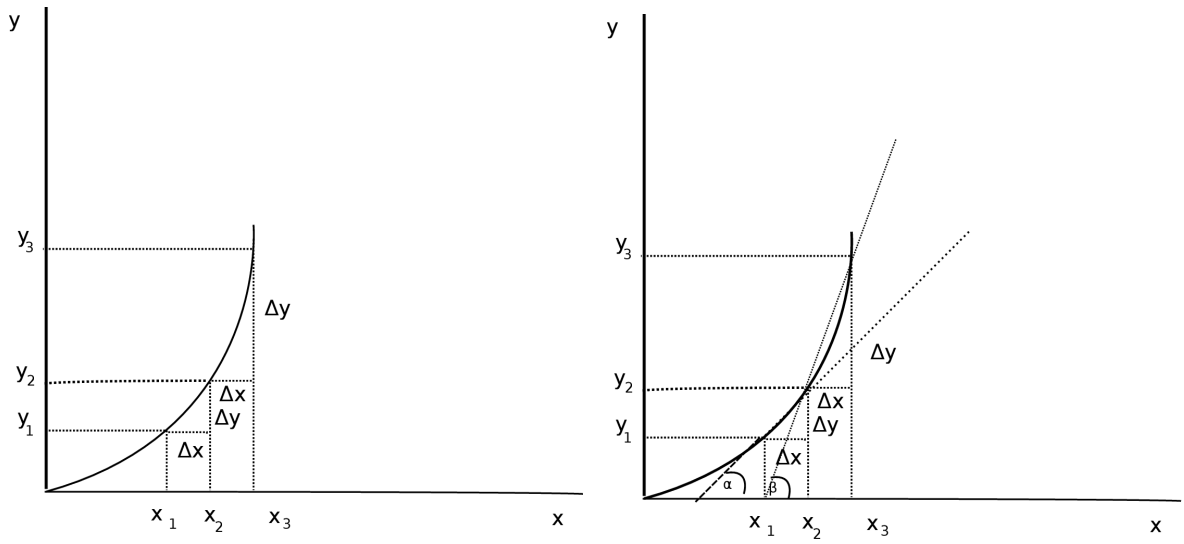


Figura 11

Il medesimo procedimento può evidentemente essere applicato per analizzare le proprietà (ed in particolare la curvatura) di funzioni decrescenti: nelle due figure che seguono sono così evidenziati i due casi, rispettivamente, di un aumento e di una diminuzione del valore assoluto del rapporto incrementale (e quindi di un aumento e di una diminuzione del valore assoluto dell'angolo che le rette formano con l'asse delle ascisse) per valori sempre più elevati della variabile indipendente.

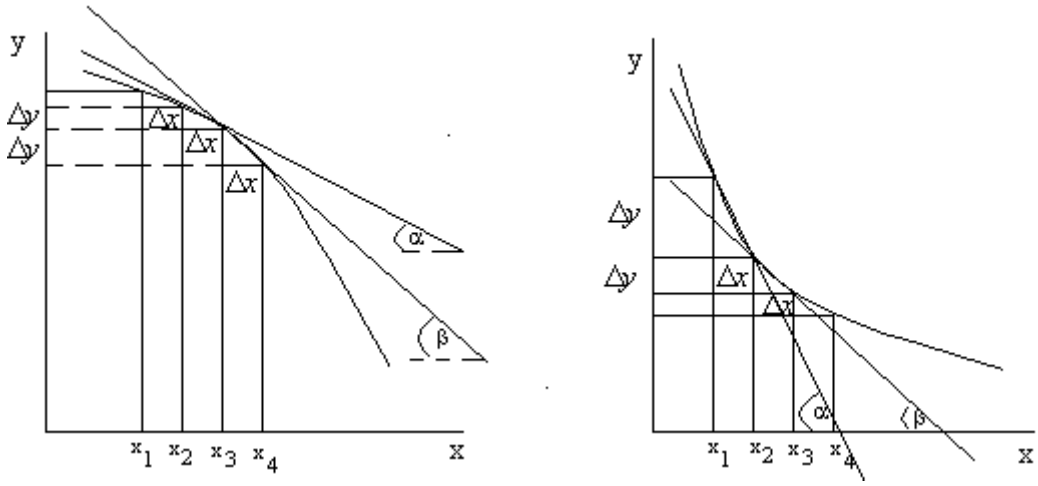


Figura 12

È ovvio che quanto più piccolo è l'incremento Δx tanto più "precisa" sarà l'informazione sulla curvatura della funzione. Ad esempio, considerando la seguente funzione:

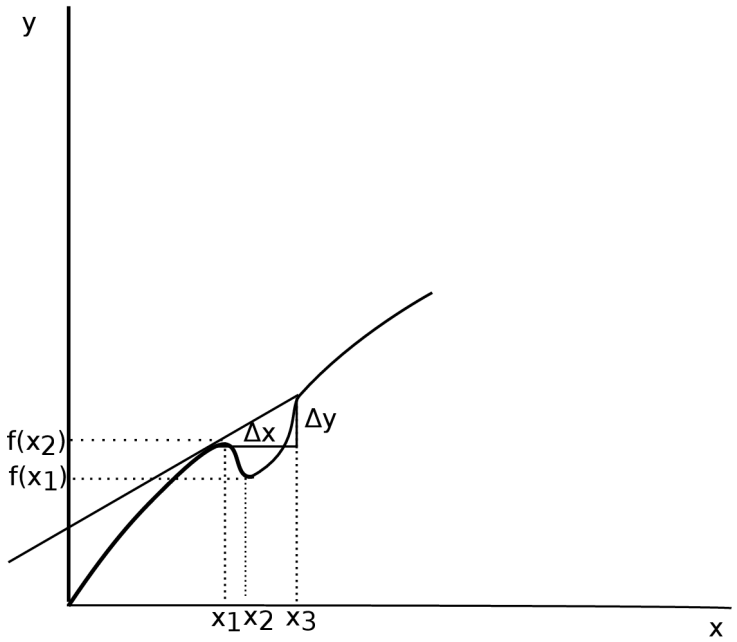


Figura 13

difficilmente, con l'incremento considerato, saremmo in grado di comprendere che $\frac{\Delta y}{\Delta x}$ è negativo nell'intervallo $x_1 - x_2$, è positivo e crescente nell'intervallo x_2-x_3, e torna ad essere positivo e decrescente per $x > x_3$.

Possiamo quindi immaginare di ridurre progressivamente l'entità dell'incremento $\Delta x \equiv x_n - x_m$ al fine di ottenere un'informazione più *puntuale* della curvatura della funzione. Tornando alla precedente Figura 10 è immediato verificare come ciò comporti un progressivo avvicinamento, sulla funzione, dei punti $f(x_n)$, $f(x_m)$ e quindi una progressiva diminuzione di Δy; se supponiamo di far tendere l'incremento Δx a valori prossimi allo zero, ovvero se supponiamo di effettuare una variazione *infinitesimale* della variabile indipendente, i punti $f(x_n)$, $f(x_m)$ tendono a diventare coincidenti e la retta che li congiunge a diventare *tangente* alla funzione, come si vede dalla seguente Figura 14 (che il lettore è invitato a confrontare con la precedente Figura 10):

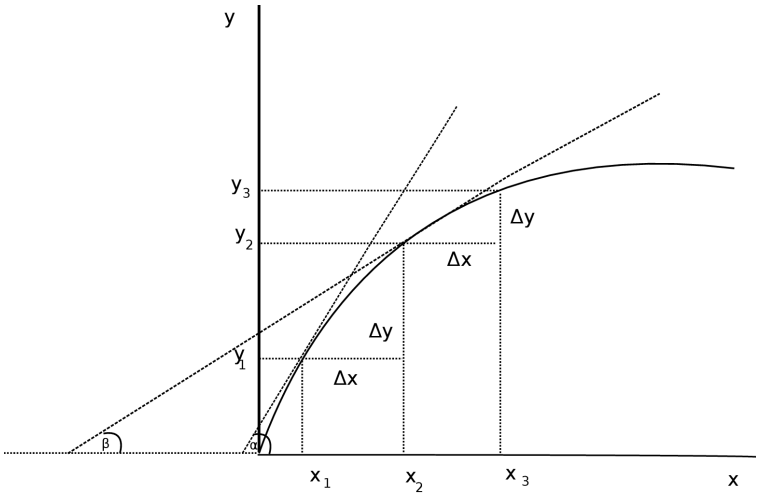


Figura 14

Siamo così in grado di comprendere, intuitivamente, la nozione di *derivata* in un punto di una funzione: data una funzione $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$, $y = f(x)$, e definito genericamente $\Delta x \equiv x - x^*$, il valore *limite* a cui tende il rapporto incrementale, al tendere di Δx a zero (e quindi al tendere di x ad x^*), *se esiste*, si chiama derivata della funzione in x^*, e si indica generalmente con $\frac{dy}{dx}$ ovvero con $f'(x^*)$[19]. La derivata di una funzione in un punto x^*, geometricamente, sarà allora espressa dal coefficiente angolare della retta *tangente* la funzione in $f(x^*)$: conformemente alle considerazioni precedenti, se tale coefficiente è crescente (decrescente) per valori sempre più elevati della x diremo che la derivata della funzione è *crescente* (decrescente) lungo il dominio della funzione ed ovviamente diremo che la derivata è *positiva* (*negativa*) se il valore del coefficiente angolare è positivo (negativo)[20]. Nella figura seguente troviamo un esempio di funzione con derivata positiva crescente:

[19] Si noti che il simbolo d (in luogo di Δ) nel rapporto incrementale indica appunto la considerazione, adesso, di una variazione *infinitesimale* (delle variabili x ed y).

[20] È evidente che nel caso della funzione affine: $y = mx + q$ ($m, x, q \in \mathbf{R}$) la derivata è *costante* ed il suo valore è espresso dal coefficiente angolare, cioè: $\frac{dy}{dx} = m$.

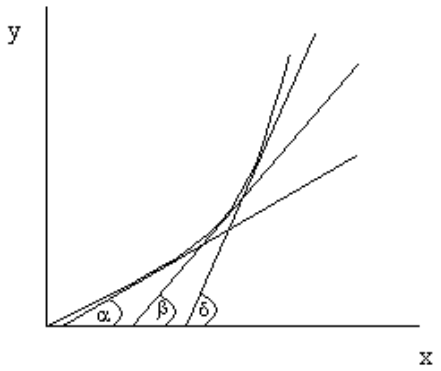


Figura 15

Da quanto detto ne consegue che *la derivata di una funzione costante è nulla* (è il caso infatti di una retta con coefficiente angolare pari a zero e quindi parallela all'asse delle ascisse: $y = q$, $q \in \mathbf{R}$, $q \neq 0$), mentre la derivata in un punto x^* del prodotto di una funzione per una costante è pari alla costante per la derivata della funzione in quel punto, ovvero:

data la funzione $f : A \rightarrow \mathbf{R}$, $A \subseteq \mathbf{R}$, $y = q\, h\,(x)$, ove $q \in \mathbf{R}$, $q \neq 0$, $\dfrac{dy}{dx} = q\, h'(x^*)$.

Quest'ultimo risultato è di dimostrazione immediata essendo il rapporto incrementale per la funzione $q\, h\,(x)$ dato da: $q\dfrac{h(x)-h(x^*)}{\Delta x}$, ovvero da $q\dfrac{\Delta h}{\Delta x}$, da cui il limite[21] di $q\dfrac{\Delta h}{\Delta x}$ per $x \rightarrow x^*$ è uguale a $qh'(x)$.

Esempio 7

Nella teoria economica molto spesso il concetto di variazione *infinitesimale* è espresso con il termine di variazione *al margine*: il *prodotto marginale* del fattore produttivo lavoro L, ad esempio, è quindi pari alla variazione del prodotto Y che si ottiene impiegando un'unità infinitesimale aggiuntiva di lavoro[22], ovvero è pari a = $\dfrac{dY}{dL}$; l'*utilità marginale* che il consumo di un bene arreca ad un consumatore è pari $\dfrac{dU}{dq}$, ove U è l'utilità e q la quantità del bene in questione ecc.[23].

[21] Rigorosamente: $\lim_{x \to x^*} q\dfrac{h(x)-h(x^*)}{x-x^*} = q\dfrac{dh}{dx} = qh'(x^*)$

[22] Si pensi qui al lavoro L misurato in termini di quantità di *ore* (*tempo*) di lavoro prestato: un'unità infinitesimale di lavoro sarebbe quindi pari ad una frazione infinitesima di un'ora di lavoro .

[23] Non deve meravigliare che i primi economisti che introdussero questo tipo di analisi, nella seconda metà dell'800, sarebbero poi divenuti famosi con il nomi di *marginalisti*.

Nello specifico la teoria della produzione assume che la frontiera tecnologica efficiente sia descrivibile attraverso la funzione di produzione $f : A \rightarrow \mathbf{R}, A \subseteq \mathbf{R}^2$, definita da: $Y = f(K, L)$, dove K è l'*input capitale*, L *l'input* lavoro e Y la *massima*[24] quantità di prodotto (*output*) ottenibile attraverso tutte le possibili combinazioni di capitale e lavoro. Nel *breve periodo*, poi, si assume che la quantità di capitale sia *data* (richiedendo del tempo l'aumento di K) e che quindi si possa incrementare il prodotto soltanto attraverso successive immissioni di lavoro: la frontiera tecnologica efficiente sarà allora espressa dalla *funzione di produzione di breve periodo*: $f : A \rightarrow \mathbf{R}, A \subseteq \mathbf{R}$ definita da: $Y = f(\overline{K}, L) = g(L)$. L'ipotesi fondamentale che poniamo su questa funzione è che il prodotto marginale del lavoro sia sempre positivo, seppure decresca continuamente all'aumentare dell'impiego di tale fattore produttivo (tendendo a zero per valori infinitamente grandi di L)[25]. Pertanto una rappresentazione geometrica conseguente di tale funzione è quella fornita nella figura che segue:

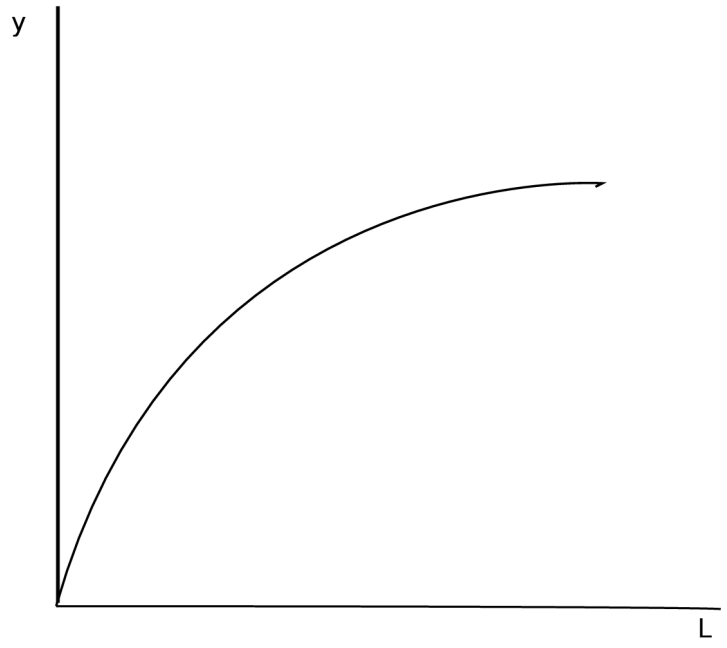


Figura 16

ove, come si può notare, $\frac{dY}{dL}$ è sempre positivo, seppure decrescente all'aumentare di N: è sufficiente infatti tracciare le rette tangenti in ogni punto alla funzione per realizzare come l'ampiezza degli angoli che queste formano con l'asse delle ascisse diminuisca continuamente all'aumentare di N.

[24] Per questo la funzione di produzione incorpora già la configurazione *efficiente* dei fattori produttivi.

[25] In realtà, più correttamente, in un primo tratto della funzione di produzione la produttività marginale dovrebbe essere assunta crescente, divenendo decrescente solo da un certo livello di impiego di L in poi (implicando una funzione di produzione a forma di *S*). Nel seguito ipotizziamo che la produttività marginale sia sempre decrescente solo a scopo esemplificativo.

Una particolare rilevanza acquista, all'interno della teoria economica, la *funzione della produttività marginale del lavoro*, ovvero *la funzione che associa ad ogni L il corrispettivo* $\frac{dY}{dL}$, e che viene evidentemente derivata dalla funzione di produzione che abbiamo appena visto: immaginando infatti di misurare gli angoli (o meglio, le tangenti trigonometriche degli angoli) che le rette tangenti alla funzione g(L) formano con l'asse delle ascisse -- per differenti valori di N -- otteniamo immediatamente (a livello grafico) la nostra nuova funzione $\frac{dY}{dL} = h$(L):

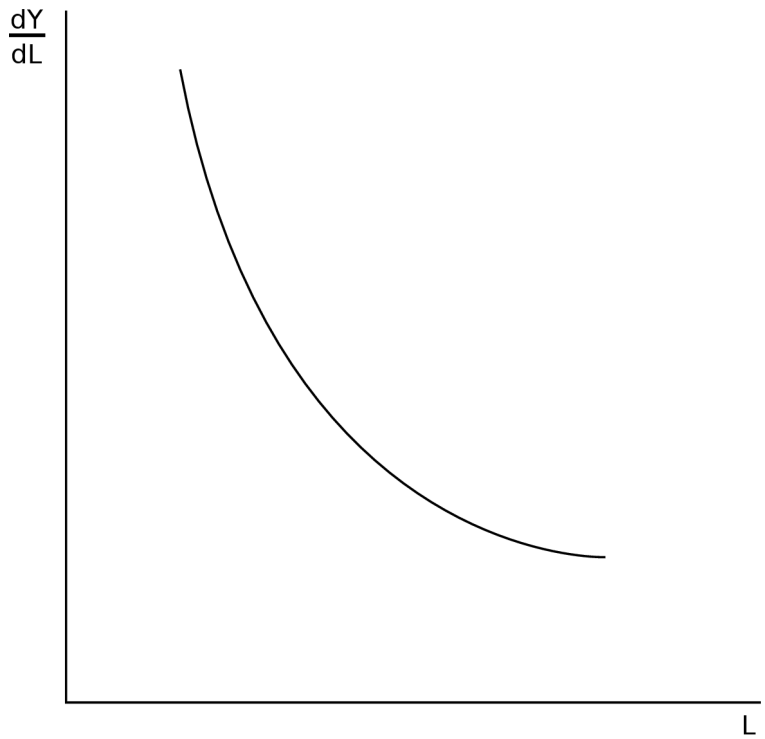


Figura 17

Esempio 8

Ritornando alla *funzione della produttività marginale del lavoro* vista nell'esempio precedente, possiamo adesso facilmente comprendere perché essa rappresenti anche la *funzione della domanda di lavoro di breve periodo* da parte dell'imprenditore. L'assunzione comportamentale fondamentale è quella che l'obiettivo imprenditoriale consista nella *massimizzazione del profitto* Π, per definizione pari alla differenza tra i ricavi totali e costi totali. Nel breve periodo, come già rimarcato, l'unico fattore produttivo variabile è rappresentato dal lavoro L: se supponiamo di restringere, per semplicità, l'analisi soltanto a quest'ultimo[26], avremo che i costi *C* saranno rappresentati dal monte salari corrisposto ai lavoratori, mentre i ricavi *R* dalla quantità di prodotto moltiplicata per il suo prezzo, ovve-

[26] Tralasciamo quindi per semplicità la quantità di capitale: questa ipotesi semplificatrice non altera le conclusioni dell'analisi che segue.

ro avremo che (ricordando che in concorrenza perfetta l'imprenditore deve assumere come dati tutti i prezzi e le remunerazioni dei fattori produttivi): Π = pY - wL, ove p = prezzo del prodotto, Y = quantità di prodotto, w = salario unitario, L = numero di lavoratori impiegato[27]. Essendo (dalla funzione di produzione di breve periodo) Y = *g* (L), possiamo così riassumere l'obiettivo imprenditoriale:

max Π = p *g* (L) - wL.

Come si può notare, la funzione dei costi *C* = wL è una retta (con coefficiente angolare pari a w) mentre quella dei ricavi non è altro che la nostra funzione di produzione moltiplicata per la costante p (nel caso in cui p = 1 coinciderebbe con tale funzione): ciò implica che la nuova funzione *R* = p *g* (L) conserverà la medesima *struttura* della funzione di produzione di breve periodo, risultando semplicemente traslata verso l'alto in ragione del valore di p.

A livello grafico la massimizzazione del profitto equivale a massimizzare la distanza tra le due funzioni *R* e *C* nella zona in cui *R* > *C* (che corrisponde, a livello numerico, a massimizzare la differenza positiva tra *R* e *C*), ed è intuitivo notare che ciò avverrà nel punto in cui le due funzioni presentano la medesima inclinazione, ovvero in corrispondenza di quel L* ove vale l'uguaglianza[28]: $p\frac{dY}{dL}=w$:

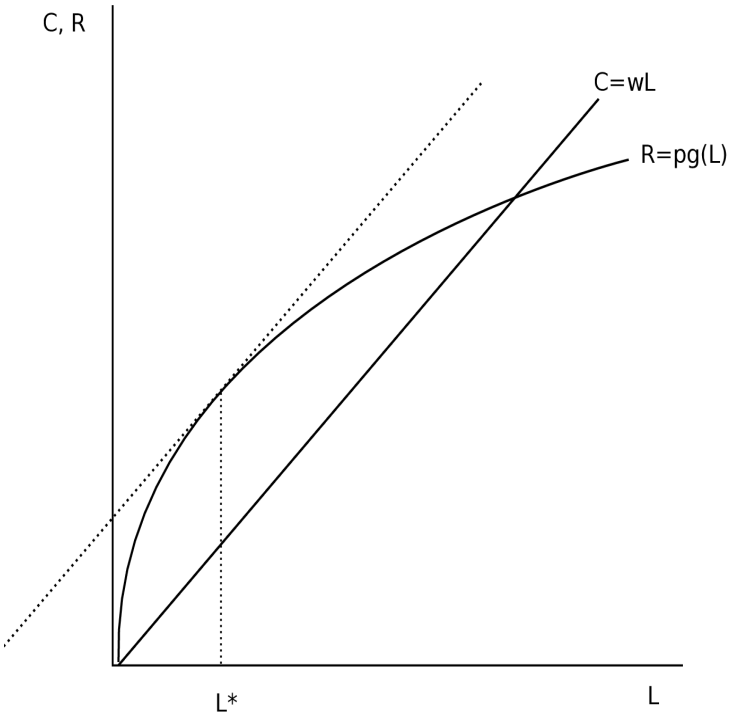


Figura 18

Abbiamo così mostrato graficamente che una condizione *necessaria* per la massimizzazione del pro-

[27] Meglio ancora (si ricordi la discussione precedente) w = salario orario; L = quantità di ore di lavoro.
[28] Si ricordi infatti che la derivata di p*g*(L) = p*g*'(L), essendo p costante.

fitto è espressa dall'uguaglianza tra il salario monetario e la *produttività marginale in valore del lavoro*, cioè: $w = p\frac{dY}{dL}$, ove appunto $p\frac{dY}{dL}$ è definita produttività marginale in valore del lavoro. Possiamo adesso fornirne una giustificazione economica: w è il costo che l'imprenditore deve sostenere per impiegare, *al margine*, un'unità aggiuntiva di lavoro, e rappresenta pertanto il *costo marginale*; $p\frac{dY}{dL}$ è invece il ricavo (espresso in termini monetari) che l'imprenditore ottiene dall'impiego di un'unità aggiuntiva di lavoro e rappresenta pertanto il *ricavo marginale*. Fintantoché quest'ultimo è superiore a w (graficamente, fintantoché la derivata della p*g*(L) è maggiore di w) l'imprenditore avrà interesse ad espandere la produzione arrestandosi soltanto quando $w=pg'(L)$ (e che si arrivi a tale punto, come si può notare graficamente, è garantito dalla forma della funzione di produzione). Se l'imprenditore infatti impiegasse lavoratori oltre il punto L*, il ricavo al margine che otterrebbe da ognuno di essi sarebbe progressivamente – e sempre più – basso del costo che dovrebbe sostenere. Da quanto detto, se ne deduce che la funzione della produttività marginale in valore rappresenta anche la *funzione di domanda di lavoro*, ovvero quella funzione $L^d = h(w)$ che associa ad ogni possibile valore del salario monetario w, la corrispondente domanda di lavoro L^d (che massimizza il profitto). Infatti, per ogni livello di w l'uguaglianza sopra scritta determina univocamente l'ammontare di lavoro domandato dall'imprenditore.

Si noti che riscrivendo la condizione di equilibrio $w = p\frac{dY}{dL}$, e dividendo entrambi i membri per p, si ottiene: $\frac{w}{p} = \frac{dY}{dL}$, da cui è possibile derivare la formulazione alternativa della domanda di lavoro come quella funzione $L^d = f\left(\frac{w}{p}\right)$ che associa ad ogni livello del *salario reale* $\frac{w}{p}$ la corrispondente domanda di lavoro L^d (che massimizza il profitto). In tale formulazione alternativa la funzione di domanda di lavoro *coincide* con la funzione della produttività marginale. In questo caso, per ogni livello di $\frac{w}{p}$ l'uguaglianza sopra scritta determina univocamente l'ammontare di lavoro domandato dall'imprenditore. Qualunque formulazione si scelga, è evidente la relazione decrescente esistente tra L^d e w: all'aumentare del salario monetario la domanda di lavoro diminuisce (e viceversa):

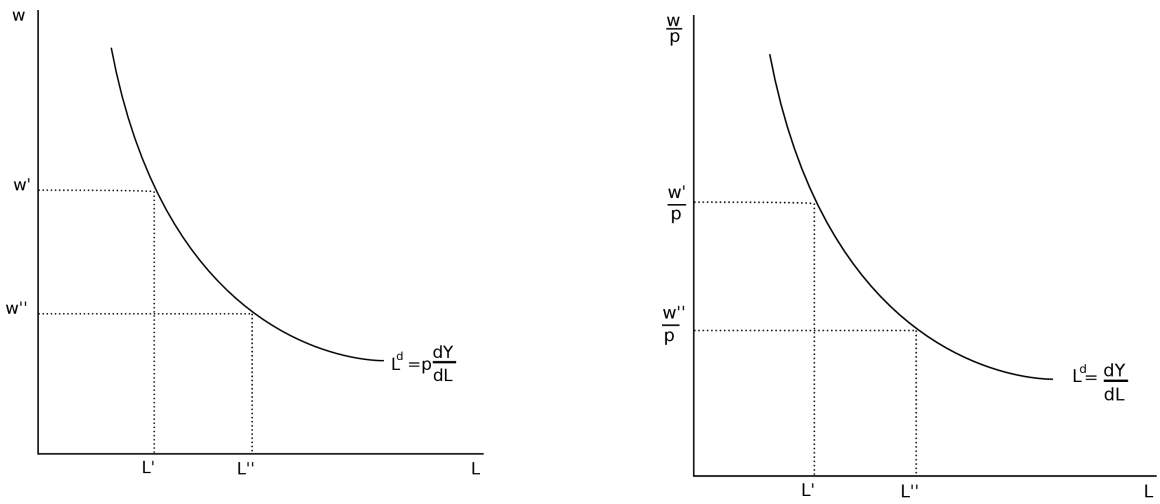


Figura 19

Un'ultima classe di funzioni elementari su cui merita brevemente soffermarsi sono le *iperboli equilatere*, ovvero quelle funzioni f : A → **R**, A ⊆ **R**, definite da: $y = \frac{ax+b}{cx+d}$ con a, b, c, d ∈ **R**, $x \neq$ -d/c, c ≠ 0[29]. Come si può notare dall'equazione riportata, al tendere della variabile indipendente al valore che annulla il denominatore la funzione non è definita, ovvero il valore della variabile dipendente tende ad infinito, cioè:

per x → - d/c, y → ± ∞.

Inoltre, se si riscrive la funzione in maniera conveniente dividendo il numeratore ed il denominatore del secondo membro per x (ovviamente $x \neq 0$), ottenendo $y = \frac{a + b/x}{c + d/x}$, si può immediatamente notare che per valori sempre più grandi in valore assoluto[30] della variabile indipendente la variabile dipendente tende al valore costante a/c, cioè:

per $|x|$ → ∞, y → a/c

I due valori notevoli che abbiamo così ottenuto individuano graficamente due rette, rispettivamente di equazione x = - d/c e y = a/c: alla prima viene dato il nome di *asintoto verticale* dell'iperbole, alla seconda quello di *asintoto orizzontale* della medesima, e praticamente sono quelle due rette che l'iperbole approccia al tendere rispettivamente di |y| → ∞ e di $|x|$ → ∞.

Esempio 9

Data l'iperbole equilatera $y = \frac{2x+7}{x+4}$, gli asintoti (orizzontale e verticale) sono: y = 2 e x = - 4 e la

[29] Si noti che le ultime due disuguaglianze discendono dall'impossibilità, nel campo dei numeri reali, di dividere un qualunque numero per zero.

[30] Si ricordi che la scrittura $|x|$ → ∞ significa x → ± ∞ .

conseguente rappresentazione grafica sarà:

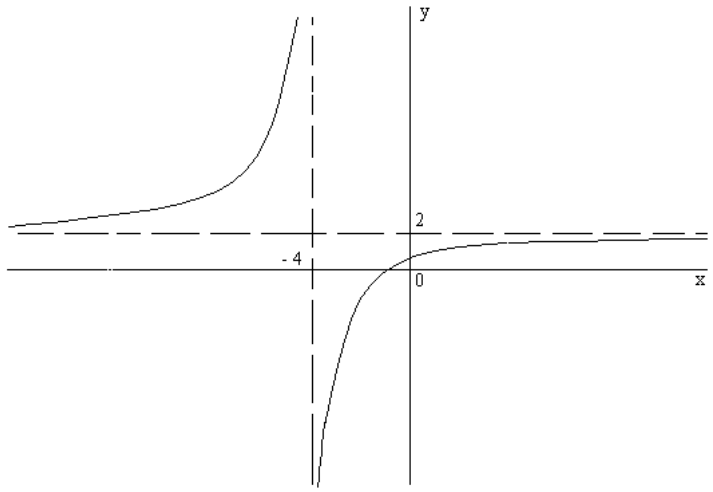


Figura 20

Un caso particolare è quello in cui a = d = 0, divenendo l'equazione $y=\frac{b}{cx}$, ovvero $xy=\overline{C}$, ove $\overline{C}\equiv b/c$ è una costante assegnata: a quest'ultima equazione dedicheremo le prossime brevi considerazioni restringendo l'analisi ai soli valori positivi delle variabili, che è quanto qui ci interessa: ciò implica assumere $\overline{C}>0$ (nel caso contrario, infatti, una delle due variabili dovrebbe essere negativa)

Per quanto concerne gli asintoti, nel caso dell'iperbole che stiamo considerando il loro calcolo è immediato: il denominatore si annulla per per $x \to 0$, mentre per $x \to \infty$, $y \to 0$; pertanto i due assi cartesiani (di equazioni rispettivamente $x = 0$ ed $y = 0$) saranno gli asintoti verticale ed orizzontale.

La precedente figura 19 ha messo in evidenza un'altra fondamentale caratteristica delle funzioni iperboliche: esse presentano due *bracci* simmetrici rispetto all'intersezione degli asintoti. Nel caso in questione ciò implica che la funzione $xy=\overline{C}$ sarà *simmetrica* rispetto all'origine degli assi cartesiani, ovvero che se la coppia (x, y) soddisfa la funzione, anche la coppia ($-x$, $-y$) soddisferà la stessa[31].

[31] E questa proprietà è ovviamente di evidenza immediata. Se ad esempio consideriamo l'iperbole $xy = 4$, le coppie (1, 4), (-1, -4); (2, 2), (-2, -2) ecc. soddisferanno tutte l'equazione scritta.

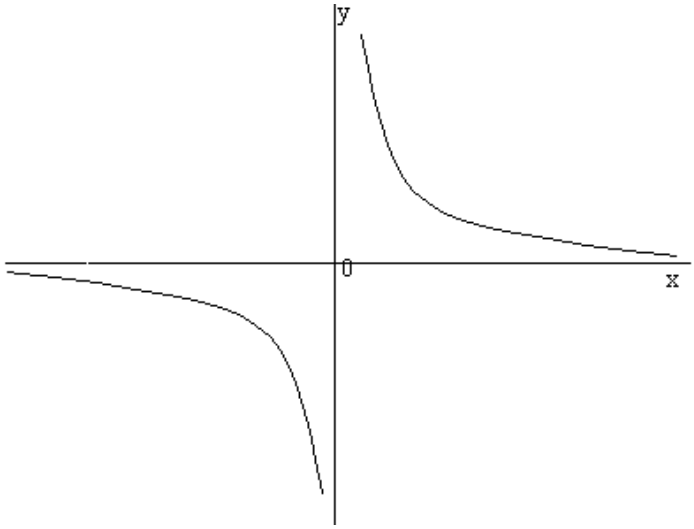


Figura 21

Per la precedente restrizione ci limiteremo pertanto a considerare il solo braccio individuato nel primo quadrante (il quadrante positivo), ovvero a considerare la funzione: $f : A \rightarrow \mathbf{R}^*$, $A \subseteq \mathbf{R}^*$ definita da $xy = \overline{C}$ (ove evidentemente, $x, y, \overline{C} \in \mathbf{R}^*$).

Un'ultima considerazione intuitiva permetterà di comprendere più a fondo le proprietà e le caratteristiche dell'iperbole equilatera che stiamo analizzando. L'equazione $xy = \overline{C}$ ci informa che, graficamente, l'iperbole descrive il luogo dei punti (x, y) il cui *prodotto* è sempre pari alla costante $\overline{C}$: ciò implica, geometricamente, che prendendo un qualunque punto sull'iperbole di coordinate (x, y), il risultante rettangolo con un vertice in questo punto e due lati sugli assi cartesiani presenterà un'area costante al variare del punto sull'iperbole[32]. Da ciò discende, intuitivamente, che variazioni di $\overline{C}$ comporteranno traslazioni della funzione: "verso l'alto a destra" per aumenti di $\overline{C}$ e "verso il basso a sinistra" per sue diminuzioni, in modo da riflettere l'aumento o la diminuzione delle aree suddette.

[32] È evidente infatti che x ed y rappresentano le lunghezze dei due diversi lati (la "base" e "l'altezza") di tale rettangolo.

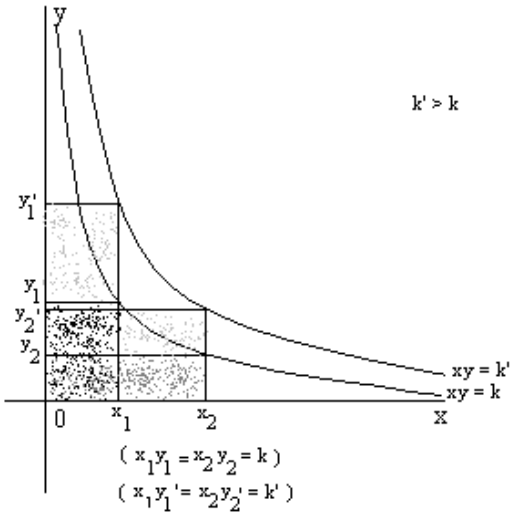


Figura 22

Esempio 10

La *teoria quantitativa della moneta* è definita dall'equazione (che è in realtà un'identità): MV = PY, ove M = quantità nominale di moneta in circolazione, V = velocità di circolazione della moneta, P = livello dei prezzi, Y = reddito (prodotto) reale. Posti M e V costanti, la precedente equazione può essere interpretata come una funzione f : A → **R***, A ⊆ **R***, definita da: $PY = \overline{C}$ (ove $\overline{C} = MV$), funzione che come si vede è esattamente un'iperbole equilatera nella forma subito sopra analizzata.

Le medesime considerazioni valgono se si intende l'equazione su scritta come una funzione di domanda di moneta (la cosiddetta *Equazione di Cambridge*): $M^d = k$ PY, con $k \equiv 1/V$. Ponendo la condizione di equilibrio M^d = M, ove M è la quantità nominale di moneta in circolazione (offerta di moneta); si ha infatti: M = k PY, che è sempre la medesima iperbole equilatera.

Conformemente a quanto visto, ogni aumento della quantità nominale di moneta in circolazione M e/o della velocità di circolazione della moneta V comporterà una traslazione verso l'alto dell'iperbole; il contrario nel caso di una diminuzione di M e/o di V.

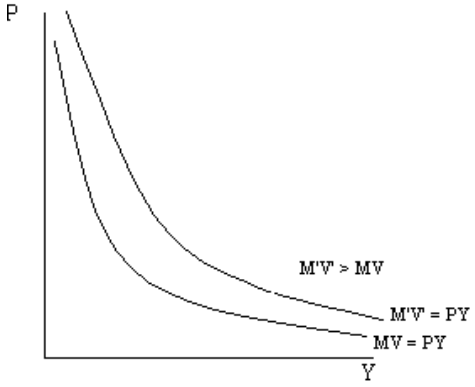


Figura 23