per un'impresa meccanica. La distinzione in oggetto è di natura sostanzialmente concettuale (e non temporale) e consente di affrontare lo studio del comportamento dell'impresa per gradi successivi di complessità.

4. LA TECNOLOGIA DI PRODUZIONE

In ogni società, in qualunque epoca storica, il patrimonio di conoscenze relative al modo in cui, a partire da certi fattori, è possibile arrivare a determinati prodotti costituisce la *tecnologia disponibile*, la quale non è altro che l'insieme delle tecniche di produzione disponibili in un dato momento.
Gli aspetti principali sui quali concentreremo ora l'attenzione sono i seguenti:

- vi sono tecniche diverse per produrre un medesimo output? In caso affermativo, quali relazioni sussistono tra le proporzioni di impiego dei vari input, vale a dire tra le varie tecniche di produzione? Come si intuisce, è qui in gioco un problema di *sostituibilità tra fattori produttivi*;
- se si aumentano (o si diminuiscono) nella medesima proporzione gli ammontari impiegati di tutti gli input, cosa ne è dell'output? Questo aumenterà (o diminuirà) nella stessa proporzione? Si tratta del problema della *scala di produzione*, problema che si pone solo in contesti di lungo periodo, dal momento che devono poter variare gli ammontari di tutti gli input (compresi quelli fissi).

Prima di procedere, richiamiamo l'importante nozione di **efficienza in senso paretiano**. Si considerino due tecniche qualsiasi a e b, appartenenti alla medesima tecnologia. Dal loro confronto può risultare che, rispetto a b, la tecnica a: *i*) implica lo stesso livello di output ma impiega quantità minori di alcuni input senza esigere quantità maggiori degli altri input; *ii*) impiega le medesime quantità di tutti gli input, ma implica una quantità maggiore di output. Diremo allora che *a è tecnicamente efficiente rispetto a b*: è questa la nozione di efficienza dovuta a Pareto. Come si nota, tale nozione rende possibile il confronto tra tecniche alternative di produzione senza dover ricorrere a misure di valore.

4.1. La funzione di produzione di breve periodo

Assumendo, per semplicità, l'esistenza di un solo output, la cui quantità prodotta indicheremo con y, *la funzione di produzione di breve periodo è l'insieme delle tecniche efficienti per produrre, nell'unità di tempo considerata, la quantità y di output utilizzando le quantità L e M degli input variabili (lavoro e materie prime), data la dimensione dell'impianto $\overline{K}$.*
In simboli:

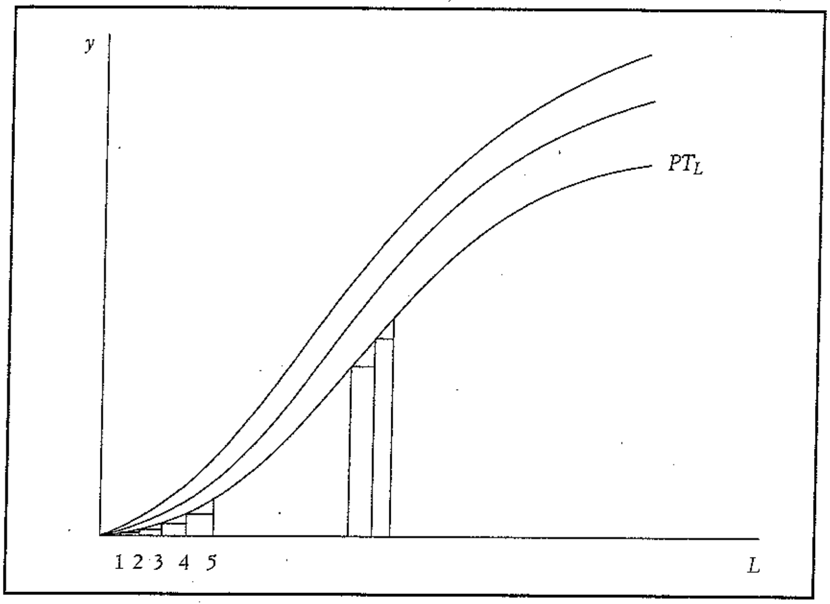


fig. 5.1. *Funzione del prodotto totale di un input variabile.*

[5.1] $$y = f(L, M; \overline{K})$$

che leggiamo: l'output y è funzione di L, M e $\overline{K}$.

Occupiamoci ora del problema *sub i)*, vale a dire della sostituibilità tra input, iniziando dapprima dal caso in cui vi sia un solo input variabile. Assumeremo cioè che L sia l'unico fattore variabile: la nostra impresa ha una quantità data di materie prime e un dato impianto; l'unico input che essa può variare al fine di ottenere quantità diverse di output è il lavoro. Sotto tale ipotesi la [5.1] può scriversi, più semplicemente,

[5.2] $$y = g\,(L)$$

per significare che y varia al variare di L, dati $\overline{M}$ e $\overline{K}$. Denominiamo la [5.2] *funzione del prodotto totale di L*. Geometricamente una delle sue possibili forme è quella di figura 5.1. La curva parte dall'origine perché, essendo L un elemento essenziale della produzione, se $L = 0$ anche $y = 0$. Chiaramente, l'andamento e la posizione di questa curva dipendono dalle quantità impiegate degli altri input. Ecco perché in figura 5.1 è disegnata una famiglia di curve del prodotto totale del lavoro. Per cogliere il significato economico della forma a sigmoide della curva PT_L è necessario introdurre i concetti di produttività marginale e di produttività media di un input.

4.2. La legge dei rendimenti decrescenti

Definiamo *produttività marginale di un input* – nel caso presente, del lavoro, che indicheremo con P_L' – la variazione dell'output dovuta ad una variazione molto piccola dell'input considerato, fermi restando gli impieghi di tutti gli altri input. In simboli:

$$P_L' = \frac{\Delta y}{\Delta L}$$

Definiamo invece *produttività media di un input variabile* il rapporto tra output ottenuto e quantità complessivamente impiegata dell'input, fermi restando gli impieghi di tutti gli altri input. In simboli:

$$PM_L = \frac{y}{L}$$

Sulla base di tali definizioni è ora agevole ricavare le curve della produttività marginale e media di un input una volta nota la curva della produttività totale dello stesso. La figura 5.2 illustra geometricamente il collegamento in questione. Come si nota, fino a che la curva PT_L è convessa (fino al punto F in fig. 5.2) la curva P_L' è crescente; dal punto in cui la PT_L inizia a diventare concava, la P_L' comincia a decrescere. Infine, allorché la PT_L raggiunge il suo punto di massimo (il punto T in fig. 5.2), la P_L' diviene nulla (la curva P_L' taglia cioè l'asse delle ascisse). Anche la curva PM_L è a forma di campana al pari della curva P_L', solo che la PM_L cresce fino al punto S della PT_L. In altri termini, il punto di massimo della curva PM_L – il punto H – è spostato a destra rispetto al punto di massimo della P_L'. Ciò riflette il fatto che la produttività media, come tutte le grandezze medie, tiene conto di tutte le unità impiegate di fattore e non solo dell'ultima unità come è invece il caso con la produttività marginale.

Osserviamo, da ultimo, che la curva PM_L raggiunge il suo massimo nel punto in cui essa è intersecata dalla curva P_L': il punto H in figura 5.2. È questa una proprietà generale che lega tra loro qualsiasi grandezza media e la corrispondente grandezza marginale: *la grandezza media consegue il suo valore massimo o minimo nel punto in cui essa eguaglia la corrispondente grandezza marginale*. Nel caso in questione la grandezza in gioco è la produttività; più avanti si tratterà di costo e di ricavo: la regola è comunque sempre la stessa.

La forma della curva PT_L di cui alla figura 5.2 cela un'importante legge economica: la **legge dei rendimenti decrescenti**. Studiando la produzione agricola, David Ricardo aveva osservato che quantità diverse di lavoro assistite da certe quantità di altri input, quali attrezzature agricole, fertilizzanti, ecc., possono essere impiegate su un dato appezzamento di terra. Nella produzio-

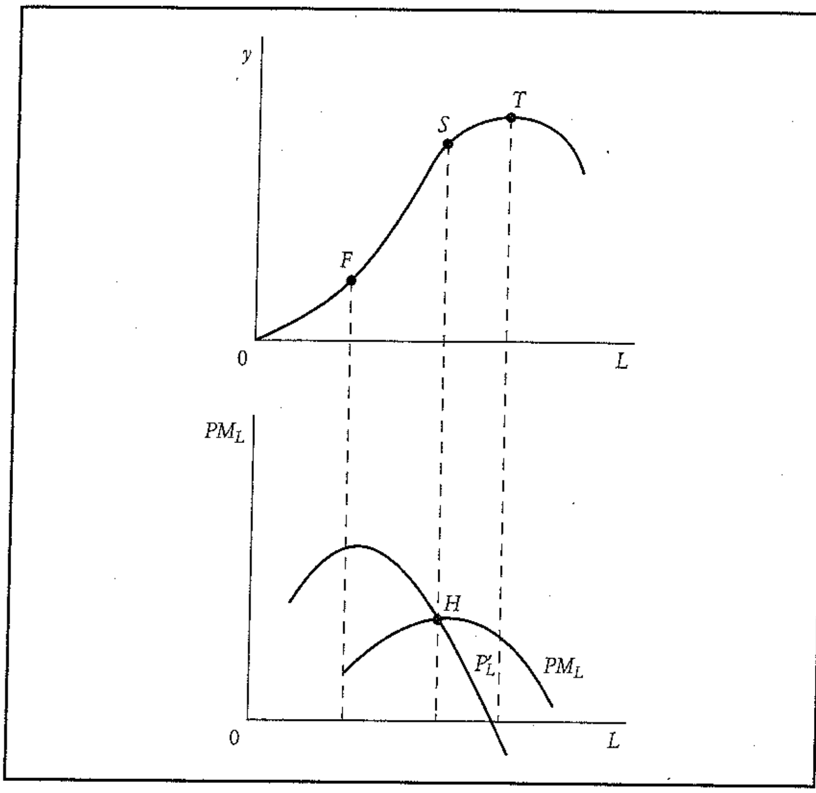


fig. 5.2. *Prodotto totale, medio e marginale di un input variabile.*

ne di beni agricoli è cioè possibile – precisava Ricardo – variare le proporzioni in cui vengono impiegati terra e «lavoro complesso» (cioè lavoro più capitale). L'economista inglese giungeva così alla legge in questione: gli aumenti di produzione risultanti da eguali incrementi nell'impiego di dosi successive di lavoro complesso, ferma restando la quantità di terra messa a coltura, prima crescono e poi decrescono. Cioè la produttività media e la produttività marginale del lavoro complesso hanno andamento prima crescente e poi decrescente, come appunto indica la figura 5.2.

4.3. Variabilità di due input

Assumiamo, ora, che due siano gli input variabili, le cui quantità indicheremo con L e M. La nostra funzione di produzione è allora:

[5.3] $$y = h\,(L, M)$$

essendo dato, come al solito, l'impianto $\overline{K}$.

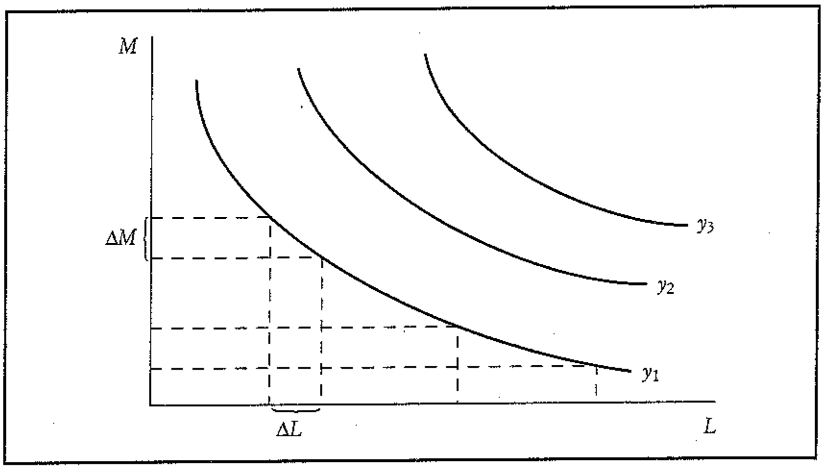


fig. 5.3. *Isoquanti di produzione.*

Per rappresentare geometricamente la [5.3] conviene servirsi di un grafico sui cui assi vengono indicate le quantità impiegate di L e di M. Otteniamo in tal modo delle curve che prendono il nome di **isoquanti di produzione** (cfr. fig. 5.3). Ciascuna curva rappresenta, infatti, tutte le combinazioni efficienti dei tassi di impiego di L e di M che danno luogo ad un medesimo livello di output y. La figura 5.3 riproduce una tipica mappa (o famiglia) di isoquanti derivata dalla funzione di produzione [5.3]. Alcune osservazioni sono qui opportune:

- poiché la funzione di produzione è, per definizione, l'insieme delle tecniche efficienti, si ha che l'isoquanto è una curva discendente da sinistra verso destra – quanto a dire che se si vuole mantenere inalterato il livello di output, all'aumento del livello di impiego di un input deve corrispondere una diminuzione della quantità impiegata dall'altro input;
- ad ogni isoquanto è associato un livello di output e questo cresce mano a mano che ci si allontana dall'origine degli assi cartesiani ($y_1 < y_2 < y_3$);
- due o più isoquanti non possono mai intersecarsi: se così fosse risulterebbe violata la proprietà di efficienza;
- gli isoquanti disegnati in figura 5.3 sono convessi rispetto all'origine. Perché? Per rispondere, si consideri il rapporto tra la variazione di M – che indichiamo con ΔM – necessaria a compensare la variazione di L – che indichiamo con ΔL – al fine di rimanere sul medesimo isoquanto, y_1. Il rapporto $\Delta M/\Delta L$ definisce il **saggio (marginale) di sostituzione** ($\text{SMS}_{M,L}$) tra M e L, vale a dire il grado di sostituibilità tra i due input ai fini dell'ottenimento di un dato volume di output.

Ora, la convessità dell'isoquanto implica che mano a mano che si procede a sostituire M con L, occorrono quantità sempre maggiori di L per «compensare» la diminuzione di una medesima quantità di M. A ben riflettere, ciò è la conseguenza diretta della legge dei rendimenti decrescenti di cui si è detto sopra. Infatti, si dimostra agevolmente che: $\text{SMS}_{M,L} = | P'_L / P'_M |$.

5. LA FUNZIONE DI PRODUZIONE DI LUNGO PERIODO: I RENDIMENTI DI SCALA

Passiamo ora al problema *sub ii*). Occupiamoci cioè dello studio delle leggi della produzione in un contesto di lungo periodo nel quale non esistono input fissi.

Con l'espressione **rendimenti di scala** ci si riferisce alla relazione tra variazione dell'output e variazioni *equiproporzionali* di tutti gli input. La qualificazione di «scala» che compare accanto al termine rendimenti indica che ciò che cambia è solo il livello di produzione, mentre la *struttura* e le caratteristiche del processo non mutano. Sia data la funzione di produzione di lungo periodo $y = f(L, M, K)$ e si assuma di partire da un livello di impiego degli input tale che l'output prodotto sia y^0. Si aumentino ora tutti e tre gli input nella medesima proporzione, poniamo di raddoppiarli, ottenendo così un nuovo livello di output $y^1 > y^0$. Ciò posto, se si trova che l'output aumenta nella medesima proporzione, cioè raddoppia, diremo che si registrano *rendimenti costanti di scala*; se l'output aumenta più del doppio, parleremo di *rendimenti crescenti di scala*; se l'output aumenta meno del doppio, parleremo di *rendimenti decrescenti di scala*.

Ma da che cosa dipendono i diversi tipi di rendimenti di scala? La causa più frequente da cui originano i rendimenti crescenti di scala è quella delle *indivisibilità tecniche*. Se è vero che un processo può sempre venire raddoppiato (o triplicato, ecc.) non sempre è possibile dimezzarne l'intensità di esercizio. Due uomini con due badili possono produrre il doppio di un solo uomo con un badile, ma un uomo a metà tempo con «mezzo badile» non produce nulla. È questo un esempio di indivisibilità tecnica.

Se i rendimenti crescenti di scala fossero sempre ed esclusivamente dovuti alle indivisibilità tecniche sarebbe possibile individuare un qualche livello di output in corrispondenza del quale tali economie di scala scomparirebbero. In altri termini, sarebbe possibile individuare una «scala ottima» di produzione. Invece, accanto alle indivisibilità tecniche troviamo, quali cause dei rendimenti di scala, le cosiddette *economie dovute alla natura tridimensionale dello spazio*. Si tratta di rendimenti di scala **statici**. Per comprendere di che si tratta si consideri la relazione area-volume. Si pensi ad un cubo di lato x. La superficie totale è pari a $6x^2$ mentre il volume è pari a x^3. Se il lato del cubo raddoppia $(2x)$ la superficie totale aumenterà di 4 volte $(24x^2)$ mentre il volume aumenterà di 8 volte $(8x^3)$. Se dunque la capacità di produzione di un impianto dipende dal volume, mentre il suo costo dipende dall'area, avremo che il costo di produzione aumenta in misura meno che proporzionale rispetto alla produzione. Si pensi ad un oleodotto: la sua capacità di trasporto può essere quadruplicata raddoppiando il suo diametro mentre il costo di costruzione è legato al diametro e non alla capacità. Altro esempio è quello di un impianto, poniamo un altoforno, che richiede quantità di input variabile (energia) per unità di output tanto più piccole quanto maggiore è la dimensione dell'impianto. Tale *space principle*

trova applicazione sia ai materiali durevoli (oleodotti) sia agli input non durevoli (contenitori di plastica) e non è legato all'indivisibilità.
Altro fattore generatore dei rendimenti crescenti di scala è rappresentato dalle *economie di specializzazione*: più elevati livelli di produzione consentono una più spinta divisione e specializzazione del lavoro con conseguenti aumenti di produttività e dunque riduzione dei costi. L'introduzione della catena di montaggio, prima, e dell'automazione, poi, e i fenomeni di *learning by doing* sono altrettanti esempi di economie di specializzazione. Come si può comprendere si tratta di veri e propri rendimenti **dinamici** di scala, ed anch'essi nulla hanno a che vedere con le indivisibilità tecniche.

5.1. Economie di scala

Con l'espressione **economie di scala** vengono designati tutti quei fattori che fanno sì che i costi unitari relativi alla produzione di livelli elevati di output siano inferiori a quelli che verrebbero sostenuti per produrre livelli più bassi del medesimo output. Ma da cosa originano le economie di scala? In relazione alla natura dei fenomeni da cui originano, si è soliti distinguere le economie di scala in *reali* e *pecuniarie*.
Economie di tipo *pecuniario* sono, ad esempio, quelle di cui beneficia l'impresa allorché riesce a pagare prezzi più bassi per gli input che utilizza, e ciò in conseguenza del fatto che al crescere della scala di produzione aumenta la quantità domandata di input da parte delle imprese. Prezzi più bassi per le materie prime; costi minori per il finanziamento esterno; minori spese di trasporto e di distribuzione dell'output, e simili, sono altrettanti esempi di economie di scala pecuniarie. Come si intuisce, si tratta di economie *esterne* all'impresa perché hanno a che vedere con la forza contrattuale delle imprese sui mercati degli input.
Le economie di scala *reali*, invece, sono quelle associate a riduzioni nella quantità impiegata degli input all'aumentare dei livelli di output di una certa impresa. Proviamo a esemplificare.
• Per iniziare una qualsiasi attività produttiva è sempre necessario un certo esborso monetario. Quanto più elevato è il livello di produzione, tanto minore sarà l'incidenza di tale esborso per unità di prodotto. Si tratta, in altri termini, di un tipico *effetto di costo fisso*. Tra i costi fissi iniziali troviamo i costi di sistemazione del macchinario; i costi per la messa a punto delle apparecchiature; i costi relativi alla linea del prodotto, ecc.
• Altro esempio di economie di scala reali è rappresentato dalle *economie di capacità di riserva*. L'impresa mantiene sempre una certa capacità di riserva e ciò allo scopo di non interrompere il flusso di output nel caso di guasti al macchinario. Ora, un'impresa di piccole dimensioni che utilizza, poniamo, una sola macchina, dovrà mantenerne due per garantirsi dal rischio di interruzioni; ma l'impresa che utilizzasse quattro macchine non avrà certo biso-

gno di otto macchine per cautelarsi: gliene basteranno sei. Lo stesso dicasi per le maestranze adibite alle riparazioni. Simile è la situazione con le *economie di scorte*; le scorte di materie prime aumentano al crescere dell'output, ma in generale l'ammontare di scorte – il cui mantenimento costituisce un costo per l'impresa – varia in misura meno che proporzionale rispetto all'aumento dell'output.

• I *rendimenti crescenti di scala*, di cui abbiamo detto sopra, sono, infine, l'altra importante categoria di fenomeni da cui originano economie di scala reali.

Un'ultima annotazione. Le economie di scala non vanno confuse con le **economie di varietà** (*economies of scope*). Quando l'impresa ha la capacità di produrre più beni – è cioè in grado di realizzare una *produzione multipla* – può accadere che, producendo più prodotti insieme, si riesca a realizzare, a parità di livelli di output, costi complessivi più bassi della somma dei costi relativi alla produzione di ciascun output a sé considerato. In questi casi si parla di economie di varietà.

5.2. Diseconomie di scala

È bene precisare che le cose fin qui dette non implicano che la grande dimensione di impresa sia sempre da preferirsi alla piccola dimensione. Può infatti accadere che in certi comparti produttivi o in certi periodi storici l'importanza che in astratto è possibile attribuire alle economie di scala non si materializzi, ovvero vengano a manifestarsi diseconomie di scala dovute a:

• limiti posti alla disponibilità di input in presenza, ad esempio, di limiti all'offerta di lavoro; l'aumento della scala di produzione può comportare costi più elevati del lavoro e ciò per la necessità di assicurarsi la manodopera necessaria;

• considerazioni meramente tecniche: all'aumento della capacità di singole unità dell'impianto può far seguito un aumento delle tensioni e/o frizioni, per cui occorrono materiali più costosi;

• problemi di management: il controllo e il coordinamento delle operazioni in impianti di grandi dimensioni possono diventare di difficile soluzione;

• fenomeni di conflittualità: è un fatto di immediata constatazione che le grandi dimensioni possano aumentare i conflitti tra lavoratori e proprietà ovvero tra proprietà e controllo dell'impresa; inoltre, la complessità delle grandi organizzazioni può talvolta determinare costi organizzativi alquanto elevati.

Spetta dunque alla ricerca empirica decidere, caso per caso, se le economie di scala, nel loro insieme, sopravanzano o meno le diseconomie di scala, tenuto conto anche del contesto storico-istituzionale nel quale si muovono l'impresa o il settore. In linea teorica non pare plausibile formulare alcuna «legge» al riguardo, anche se un minimo di riflessione ci porterebbe a concludere che in settori quali quello chimico (di base) e siderurgico è lecito attendersi la prevalenza di economie di scala.

6. IL COSTO DI PRODUZIONE

L'analisi fin qui svolta ha riguardato solo il sistema dei vincoli di natura tecnica sotto cui ha luogo l'attività produttiva. Nulla si è ancora detto dei vincoli di natura economica, tanto che, al momento, non siamo in grado di indicare quale tecnica su un dato isoquanto l'impresa sceglierà per produrre un dato livello di output. In altri termini, non disponiamo ancora di un criterio di scelta tra tecniche efficienti alternative. Per risolvere questo tipo di problemi dobbiamo introdurre una nuova categoria economica, quella di costo di produzione.

In prima approssimazione, il *costo di produzione* è l'esborso monetario che viene sostenuto per acquisire tutte le risorse necessarie all'esercizio dell'attività produttiva. In quel che segue, si assumerà che: *i*) tutti gli input vengano acquistati dall'impresa a prezzi che non dipendono dagli ammontari d'acquisto; *ii*) tutti gli input possano essere acquistati dall'impresa negli ammontari richiesti. Tali assunti implicano che i mercati dei fattori produttivi siano di concorrenza perfetta.

Infine, l'ipotesi di comportamento dell'impresa capitalistica relativa al costo di produzione è la seguente: se vi sono parecchie tecniche per ottenere un dato livello di output, ma a costi diversi, l'impresa sceglierà sempre quella che implica il minor costo – vale a dire, *l'impresa sceglie la tecnica di produzione che minimizza il costo*. Ne deriva che il costo di produzione di un certo livello di output, dati i prezzi degli input e data la tecnologia, è il *minor costo* che si riesce a realizzare per produrre quel livello di output.

6.1. Tipologia dei costi di produzione

Come si ricorderà, nel breve periodo l'output è funzione dei livelli di impiego dei fattori variabili e dei servizi forniti dall'impianto a disposizione dell'impresa. Il costo di produzione dipende perciò sia dal costo d'impiego dei fattori variabili, sia dal costo per acquisire la disponibilità dell'impianto. Chiamiamo i primi *costi variabili*, dato che essi variano al variare del livello di output. Essi includono i costi delle materie prime consumate nel corso del processo produttivo; i costi di alimentazione dei macchinari; i costi di lavoro diretto (gli addetti alla produzione); la perdita di valore delle attrezzature dovuta alla loro utilizzazione nel processo produttivo.

Chiamiamo, invece, *costi fissi* quelli che l'impresa sosterrebbe anche se il tasso di produzione fosse nullo, dal momento che si tratta di costi relativi a decisioni prese in passato, al momento cioè in cui fu deciso di acquisire l'im-

pianto. Tra i costi fissi troviamo i costi di manutenzione non connessi all'utilizzo dell'impianto; i costi di assicurazione; i costi generali di amministrazione; la remunerazione per il lavoro imprenditoriale e per il capitale finanziario investito nell'impianto.
Quest'ultima componente di costo, detta anche **profitto normale**, ha la natura di un costo opportunità: un operatore decide di realizzare l'investimento finanziario che gli consente di inserirsi in un particolare ramo d'attività in quanto prevede di ottenere un «profitto normale» almeno pari a quello che egli avrebbe potuto conseguire dirottando altrove il suo investimento finanziario.
Da quanto precede si trae che, nel breve periodo, avremo costi variabili e costi fissi. Indicando con CT i costi totali di breve periodo, scriveremo:

$$CT = CFT + CVT \tag{5.4}$$

dove CFT indica i costi fissi totali e CVT i costi variabili totali.
Rapportando i costi totali alla quantità prodotta otteniamo il **costo medio totale**, CMT. Dalla [5.4], dividendo ambo i lati per y, si ottiene:

$$CMT = \frac{CT}{y} = \frac{CFT}{y} + \frac{CVT}{y} \tag{5.5}$$

Il costo medio totale è dunque pari alla somma del costo medio fisso $CMF = CFT/y$ e del costo medio variabile, $CMV = CVT/y$.
Infine, definiamo **costo marginale** e lo indichiamo con C', la variazione del costo totale dovuta ad una variazione molto piccola del livello di output. In simboli:

$$C' = \frac{\Delta CT}{\Delta y} \tag{5.6}$$

Chiaramente, il costo marginale tiene solo conto della componente variabile del costo totale, dato che la variazione del costo fisso è, per definizione, nulla nel breve periodo.
Passando al lungo periodo, l'unico punto da tenere presente è che qui cade la distinzione fra costi fissi e costi variabili, e ciò per la semplice ragione che nel lungo periodo tutti gli input sono variabili, ivi compresi gli impianti. Tra tutti gli impianti che la tecnologia pone a disposizione dell'impresa, questa ne sceglierà uno in particolare. Una volta operata tale scelta, l'impresa viene ad operare in una situazione di breve periodo.
Consegue da ciò che sostanzialmente diversa è la rilevanza dei costi di breve e di lungo periodo: la conoscenza dei primi serve per le decisioni riguardanti la quantità da produrre; la conoscenza dei secondi serve per le decisioni di investimento.

6.2. Curve di costo di breve periodo

Consideriamo dapprima il caso in cui vi è un solo input variabile: sia esso il lavoro, L. I costi variabili si riducono perciò ai pagamenti sostenuti per l'acquisizione di tale input. In tale circostanza, le curve di costo si ricavano direttamente dalle corrispondenti curve di produttività, una volta noto il prezzo dell'input variabile. Cioè, dalla curva del prodotto medio del lavoro si trae la curva del costo variabile medio e da quella della produttività marginale del lavoro si ricava la curva del costo marginale.
Infatti, quando a variare è il solo lavoro, il cui prezzo unitario è il saggio di salario che indichiamo con w, il CMV sarà:

$$CMV = \frac{CVT}{y} = \frac{wL}{y}$$

e poiché la produttività media del lavoro è data da y/L, dalle eguaglianze precedenti si trae: $CMV = w/PM_L$. Dunque, il costo medio variabile è eguale al rapporto tra il prezzo dell'unico input variabile e la sua produttività media. Tale rapporto è anche noto come costo del lavoro per unità di prodotto (CLUP).
Per quanto concerne il costo marginale si avrà:

$$C' = \frac{\Delta CT}{\Delta y} = \frac{w\Delta L}{\Delta y} = \frac{w}{P'_L}$$

dal momento che l'incremento del costo totale dovuto ad un incremento di produzione eguaglia l'ammontare speso per l'acquisto delle unità aggiuntive dell'input variabile. Il costo marginale è dunque eguale al rapporto tra il prezzo dell'input variabile e la sua produttività marginale.
Da quanto precede si trae allora che, una volta note le curve della produttività media e marginale di un input, da esse si ricavano subito le curve del costo medio variabile e del costo marginale. La figura 5.4 è tratta dalla figura 5.2. Merita attenzione la perfetta corrispondenza tra le curve delle due figure: dove la curva della produttività marginale raggiunge il suo punto di massimo, quella del costo marginale è al suo punto di minimo. Lo stesso dicasi per le curve medie. Inoltre, quando una curva di produttività è crescente la corrispondente curva di costo è decrescente e viceversa. Infine, la curva del costo marginale interseca quella del costo medio variabile nel suo punto di minimo, così come la curva della produttività marginale interseca quella della produttività media nel suo punto di massimo.
E i costi fissi? Il costo fisso totale è, per definizione, costante nel breve periodo. Pertanto, la curva del costo medio fisso sarà un'iperbole equilatera come appare in figura 5.4. Sommando verticalmente le due curve CMF e CMV si ottiene la CMT.

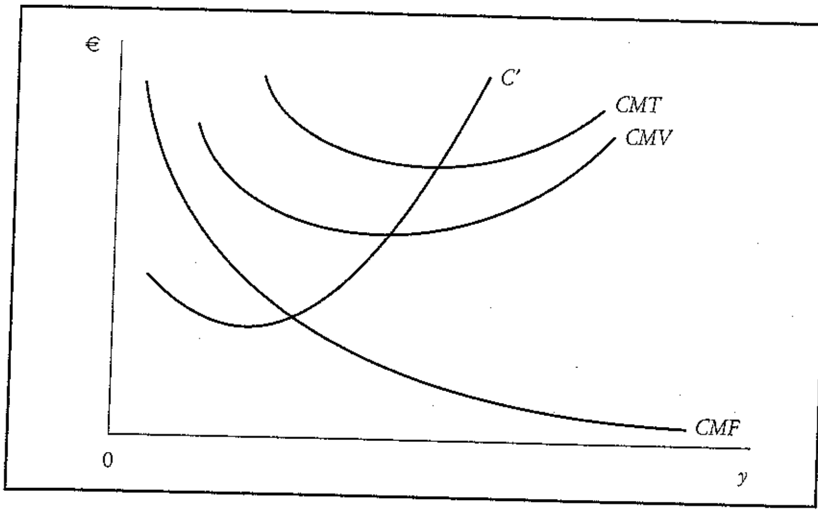


fig. 5.4. *Curve di costo medio e marginale.*

Un'ultima osservazione importante. Come si sarà notato in figura 5.4, le curve del costo marginale, del costo medio totale e del costo medio variabile hanno la forma ad *U*. Tale forma riflette la legge dei rendimenti decrescenti. Con un impianto dato, si registra dapprima una fase di produttività crescente dell'input variabile e perciò di costi decrescenti, seguita da una fase di produttività decrescente e perciò di costi crescenti. Tra queste due fasi vi è un punto in corrispondenza del quale i costi medi sono al loro livello minimo. Quando questo punto viene raggiunto, *l'impianto è utilizzato in modo ottimale*.
Per cogliere il significato di tale espressione si consideri che l'impianto è stato progettato e costruito per funzionare in combinazione con una certa quantità di input variabili. Se l'impianto viene impiegato per ottenere un livello di output minore di quello per il quale è stato progettato, non è possibile realizzare quella divisione del lavoro che l'impianto rende possibile, col risultato che gli input impiegati risulteranno meno produttivi di quel che potrebbero essere. Allo stesso modo, se l'impianto viene sovrautilizzato, si registreranno forme di congestionamento, con conseguente abbassamento del livello di produttività e quindi aumento dei costi. Vi è dunque, per ogni dato impianto, un livello dell'output in corrispondenza del quale il costo medio è minimo.

6.3. Variabilità di due input: la scelta della combinazione ottima dei fattori

Quando due sono gli input variabili, la determinazione delle curve di costo pone un problema nuovo. Infatti, come si è detto, le curve di costo indicano il livello minimo di costo per produrre un dato output. Se due sono gli input variabili, allora è necessario stabilire il criterio in base al quale l'impresa procede a realizzare la combinazione ottima dei fattori variabili. Si

tratta cioè di decidere quale punto sull'isoquanto di produzione l'impresa andrà a scegliere.
Per sciogliere l'interrogativo, è opportuno introdurre la nozione di **linea di isocosto**, intesa come luogo dei punti che indicano le combinazioni delle quantità impiegate dei due fattori variabili che implicano il medesimo costo. Indicando con w il costo unitario del lavoro e con r il costo unitario dell'altro input variabile, M, l'equazione della linea di isocosto può scriversi:

$$CVT = wL + rM$$

In figura 5.5 è rappresentata una generica linea di isocosto, la cui pendenza è data dal rapporto w/r e la cui distanza dall'origine è misurata da CVT, il costo variabile totale. Per l'isocosto valgono le stesse proprietà matematiche della linea di bilancio introdotta nel capitolo 4 a proposito del problema di scelta del consumatore: non è perciò il caso di ripeterci.
Ciò posto, la soluzione del problema riguardante la scelta della combinazione di input – e dunque della tecnica – che consente di ottenere un dato output al costo minimo può essere convenientemente espressa per via geometrica. In figura 5.5 sono rappresentati una famiglia di linee di isocosto e l'isoquanto relativo al livello prescelto di output, $\bar{y}$. Come si constata, la soluzione cercata è nel punto di tangenza E tra l'isoquanto $\bar{y}$ e una delle linee di isocosto, la CVT_2. Ogni altro punto diverso da E, comporterebbe infatti un esborso superiore per produrre $\bar{y}$ o, come nel caso di CVT_1, non permetterebbe l'ottenimento di $\bar{y}$.
Ora, poiché la pendenza dell'isocosto è data dal rapporto dei prezzi degli input, w/r, e la pendenza in un punto dell'isoquanto altro non è che il saggio marginale di sostituzione tra i due input, SMS, nel punto E di tangenza dovrà valere la seguente relazione:

$$\text{SMS} = \frac{w}{r}$$

D'altro canto, poiché il SMS è eguale al rapporto tra le produttività marginali dei due input, si ha in E:

[5.7a] $$\frac{P'_L}{P'_M} = \frac{w}{r}$$

ovvero

[5.7b] $$\frac{P'_L}{w} = \frac{P'_M}{r}$$

In definitiva, tutti i punti situati sull'isoquanto denotano, come sappiamo, tecniche efficienti; tuttavia, solo la tecnica E è quella *ottimale*, quella cioè che

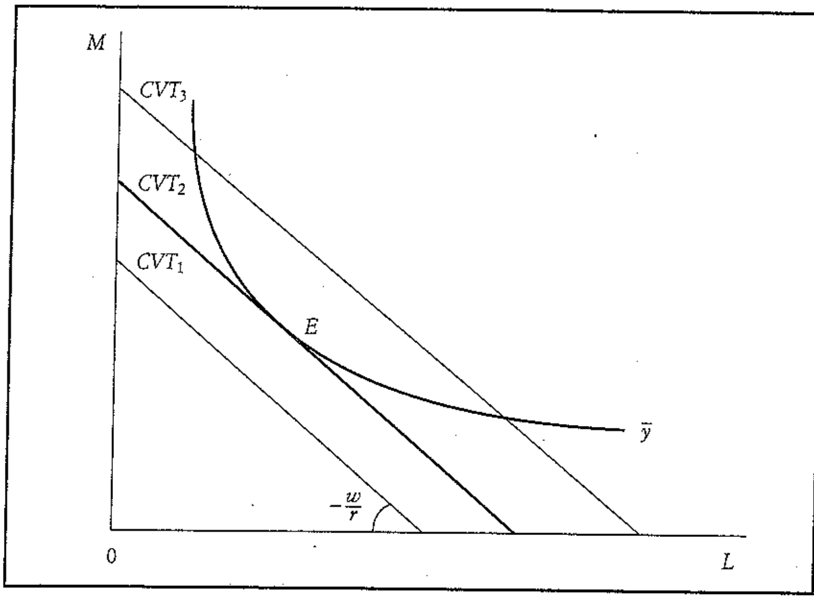
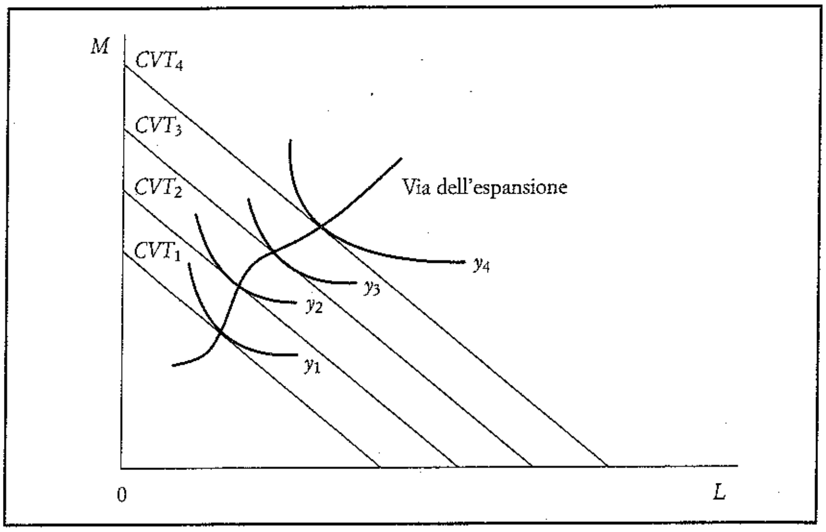


fig. 5.5. *Minimizzazione del costo totale e scelta della tecnica ottimale.*

minimizza il costo di produzione. Si conclude allora che la *tecnica ottimale*, cioè *la combinazione ottimale dei fattori, è quella per la quale il rapporto tra le produttività marginali dei due input eguaglia il rapporto dei loro prezzi*, ovvero *quella che realizza l'eguaglianza delle produttività marginali ponderate degli input*. La regola ora individuata ci permette anche di analizzare gli effetti connessi a variazioni nel prezzo degli input. Una variazione non equiproporzionale dei prezzi dei due input provoca una sostituzione della tecnica produttiva: l'impresa passa a sostituire la tecnica «più costosa» con quella «meno costosa». In quale direzione ha luogo tale sostituzione? La sostituzione tra input procede nel senso che *viene impiegato più intensamente l'input il cui prezzo relativo è diminuito e viceversa*. Così, se w diminuisce, fermo restando r, all'impresa converrà impiegare più lavoro e meno M: ciò trova chiara conferma nella relazione [5.7]. Infatti, se w diminuisce deve diminuire P'_L, il che comporta un uso più intensivo di L rispetto a M.

Siamo ora in grado di dare soluzione al problema di partenza. Per arrivare alla curva del costo totale variabile basterà rappresentare sul medesimo grafico una famiglia di isoquanti e una famiglia di isocosti: si cercherà poi il punto di tangenza di ciascun isoquanto con una linea di isocosto (cfr. fig. 5.6). Unendo questi punti di tangenza si ottiene la *via dell'espansione dell'impresa*, una curva, cioè, i cui punti indicano il costo minimo per produrre livelli successivi di output. Ma questo è proprio il significato della curva del costo totale variabile. Così, per produrre y_1 il costo totale variabile minimo è CVT_1; per produrre y_2 il costo totale variabile minimo è CVT_2 e così via. La curva CVT, dunque, si ricava dalla via dell'espansione dell'impresa. Che forma avrà la curva dei costi totali variabili? Si dimostra che anche nel caso di variabilità di due

fig. 5.6. *La via dell'espansione dell'impresa.*



input la relazione fra costo marginale e produttività marginale degli stessi è la medesima di quella che abbiamo visto sussistere nel caso in cui un solo input è variabile. Si conclude pertanto che anche nel caso presente, le curve di breve periodo dei costi medi e del costo marginale assumono la tipica forma ad *U* già vista nel caso precedente.

6.4. Curve di costo di lungo periodo

Nel lungo periodo, l'impresa può scegliere, come sappiamo, il tasso di impiego di tutti gli input; non vi sono perciò input fissi la cui disponibilità ponga vincoli alla capacità produttiva dell'impresa. In una situazione di lungo periodo, come procede l'impresa alla scelta dell'impianto che più le conviene? Il problema nasce dalla circostanza che associato ad ogni tipo di impianto vi è un livello di output che può essere ottenuto al costo medio minimo e quanto maggiore è la dimensione dell'impianto, tanto maggiore risulterà, in generale, il livello di output che è possibile produrre al costo medio minimo. Ciò in quanto maggiore la dimensione d'impianto, più elevato il costo fisso e quindi più elevata l'incidenza del costo fisso sul costo medio per bassi livelli di produzione.
Per chiarire, si assuma che l'impresa possa scegliere tra tre impianti di diversa dimensione: piccola, media, grande. L'impianto di piccola dimensione implichi una curva dei costi medi quale la CM_1 della figura 5.7*a*, quello di media dimensione la CM_2 e quello di grande dimensione la CM_3. Se l'impresa prevede di vendere, e dunque programma, un livello di output non superiore a y_1, essa sceglierà certamente l'impianto piccolo; se invece prevede di vendere un livello di output compreso tra y_1 e y_2 essa opterà per l'impianto di me-

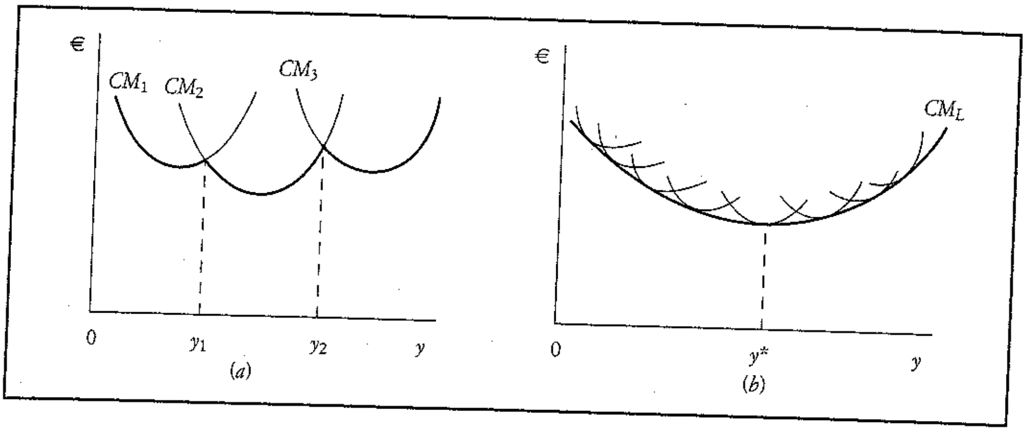


fig.5.7. *Curva di costo di lungo periodo.*

dia dimensione; se infine si aspetta di poter vendere un livello di output superiore a y_2 essa deciderà per l'impianto di grande dimensione.

Se l'impresa, dopo aver optato per l'impianto di piccola dimensione, si trova a dover produrre un livello di output maggiore di y_1, essa potrà ottenerlo ad un costo superiore rispetto a quello che avrebbe sostenuto se avesse inizialmente acquisito l'impianto di media dimensione, ecc. Sulla falsariga di tale ragionamento si conclude che la curva segnata in neretto in figura 5.7*a* non è altro che la *curva del costo medio di lungo periodo*, nel caso di esistenza di tre soli impianti.

Se ora assumiamo che gli impianti disponibili, anziché tre, siano in numero infinito, la curva spezzata della figura 5.7*a* diviene una curva liscia (cfr. fig. 5.7*b*). Tecnicamente, essa è l'**inviluppo** delle curve di costo di breve periodo e denota appunto la curva del costo medio di lungo periodo: la CM_L. Perché anche la curva CM_L ha la forma ad *U*, come le curve del costo medio di breve periodo? La risposta che viene data dalla teoria neoclassica è che ciò dipende dalla legge dei rendimenti di scala: fino ad un certo livello di output y^* il costo medio diminuisce all'aumentare della dimensione dell'impianto per il prevalere dei rendimenti crescenti di scala; superato quel livello di output, il costo medio incomincia a salire all'aumentare della dimensione dell'impianto a seguito del prevalere dei rendimenti decrescenti di scala, dovuti soprattutto a fattori di natura organizzativa.

Da quanto precede si trae che la natura della curva del costo medio di lungo periodo è quella di una *planning curve*, di una curva cioè sulla cui base l'impresa decide quale impianto porre in esercizio al fine di produrre al costo più basso il *livello previsto* di output. E poiché si assume che il numero degli impianti disponibili sia infinito, a ciascun livello previsto di output corrisponderà uno ed un solo impianto. Non vi è alcuna flessibilità per l'impianto, nessuna capacità produttiva di riserva per far fronte ad eventuali variazioni stagionali della domanda. Ma si osservi come tutto ciò sia strettamente collegato all'*assunto di previsione perfetta* dei livelli di output da produrre da parte dell'impresa. In caso contrario, infatti, l'impresa non potrebbe non riservarsi un certo margine di flessibilità.