The Production Function 2

2.1 Introduction

Economic theory is, to a large extent, about money – about costs, prices, markets, return on investment, profit and similar economic concepts. This is also the case with the theory of production economics. However, the theory of production economics is special in that the limits of economic behaviour are defined by the *technical* production possibilities. Production technology is the decisive factor regarding the quantity produced and how it may be produced. Therefore, a very important part of the theory of production economics consists of describing the production technology which defines the framework for the economic behaviour.

This chapter is concerned with the description of production technology, which is traditionally based on the production function. Apart from the production function, the chapter also introduces a number of other concepts related to the description of production technology.

2.2 Production Technology

Production technology is, in its most general form, a description of the relationship between input and produced output. The description of production technical relationships is based on empirical observation of relationships between inputs and outputs, as e.g. described in Table 2.1 which shows the relationship between the addition of nitrogen fertiliser (N) and the cereal yield.

The specified relationship can be illustrated graphically, as shown in Fig. 2.1 next to the table. *It is this curve, as shown in Fig. 2.1 that was first referred to as a production function.*

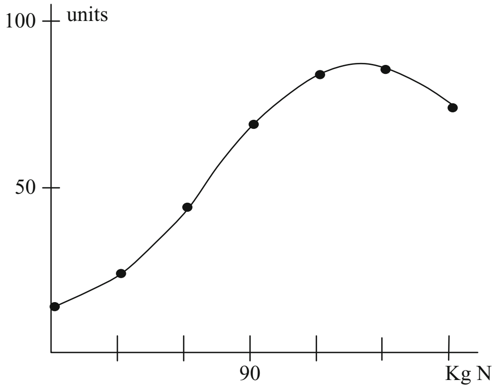
Later, in line with the development of mathematical and statistical tools for the description of production technical and economic relationships, the production function was described by means of mathematical function relationships. The choice of the functional form to illustrate the empirically observed relationships

S. Rasmussen, *Production Economics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-642-30200-8_2, © Springer-Verlag Berlin Heidelberg 2013

Table 2.1 Yield with increased N addition

Kg of nitrogen (N) per hectare	Cereal yield, units per hectare
0	15
30	25
60	45
90	70
120	85
150	85
180	75

Fig. 2.1 Production function



as a nice curve which would pass through the observed points, as shown in Fig. 2.1, and the subsequent estimation of the parameters of the function itself, came to be an important discipline in production economics.

However, to describe production technology based *solely* on observations of relationships between inputs and outputs, as shown in Fig. 2.1, is inadequate.

Firstly, it should be noted that the curve in Fig. 2.1 only describes the quantity of produced output as the function of *one* input. However, what about the other inputs used in the production? Apart from nitrogen fertiliser, the use of labour, seeds, fuel, machinery etc. is also required when growing cereal crops. Generally, production always includes at least two, and often more, inputs. A complete description of the production technology for a given product will therefore presuppose a multi-dimensional illustration providing a *simultaneous illustration* of the relationship between output and *all* inputs. With a certain level of drawing skill, such a graphical illustration is possible for productions with only two inputs. However, this is not possible if there are three or more inputs. The solution could be to describe the production technology as partial production functions, i.e. functions with only one variable input, while the remaining ones are presumed to be fixed at a given level. With e.g. eight inputs, this would require that the production technology should be illustrated as eight figures similar to Fig. 2.1. Such an illustration is, however, insufficient since the interaction between the various inputs is unclear from these partial figures.

Secondly, it is not possible to be certain that the described relationships between inputs and outputs, as shown in Table 2.1, constitute a *complete* description of the production technology. Can one, for instance, be certain that there are no other ways to produce 45 units of cereal crops than by the exact application of 60 kg of nitrogen fertiliser? What if the observations in Table 2.1 originate from a producer, who is not technically efficient, i.e. produces less for a given input level than that which is technically possible? In such a case, there would be other possible points above the curve in Fig. 2.1, which should therefore also be included to give a complete description of the production technology. The same would be true for the points below the curve. For example, is it not technically possible to produce 45 units of cereal crops through the use of 90 kg of nitrogen? Thus, the points below the curve should also be included to give a complete description of the production technology.

This shows that the act of describing the production function solely as a curve interlinking empirical observations of relationships between inputs and outputs may be much too *incomplete* and too *imprecise* a description of the production technology for a given product. The correct approach must be to describe the production technology as the complete set of all the actual possibilities at the producer's disposal.

However, how can the complete set be described in a precise and unambiguous way? How can a production technology be described in a way which leaves all possibilities open to the producer to put his/her production together in a way that is optimal for the person in question? And furthermore, how can the production technology be described in a way that makes it possible to explain empirical observations which are outside the production function in Fig. 2.1?

The strictly general point of reference would be to describe the actual possible combinations of inputs and outputs. If this set is called T, then T can be defined as:

$$\mathrm{T}(x, y) \equiv \{(x, y) : x \text{ can produce } y\} \tag{2.1}$$

in which T is the *technology set*, x is the amount of input and y is the amount of output. In this strictly general formulation, both x and y could be scalars or vectors. However, for now, both x and y should be considered as scalars (one input and one output, as in Fig. 2.1).

Looking at the production as described in Table 2.1 and Fig. 2.1, it is evident that the points $(x, y) = (0, 15), (30, 25), (60, 45), (90, 70), (120, 85), (150, 85)$, and $(180, 75)$ all belong to T as it has in fact been observed that, for these combinations of x and y, x can produce y. Furthermore, the individual points in Fig. 2.1 are connected as a smooth curve indicating that these intermediate points are also possible and therefore belong to T. By doing this, it is presumed that x can be applied in any amount (x is infinitely divisible) and that the actual observations between the already plotted points will be distributed on an even curve through the points.

However, are there other points in Fig. 2.1 that belong to T? Yes, if it is possible to produce 70 units of cereal crops with 90 kg of nitrogen (which it is according to Table 2.1), then it ought also to be possible to produce less – e.g. 45 units of cereal crops – with 90 kg of nitrogen. The reason is that under all circumstances it is possible to take the 90 kg of nitrogen and dispose of the 30 kg so that the amount actually added would be 60 kg. And with 60 kg it would of course be possible to produce 45 units of cereal crops, according to the table. A more realistic description would be to imagine an inefficient producer who, even with an addition of 90 kg of nitrogen, only achieves a yield of 45 units, exactly because the producer does not produce efficiently.

In a similar way it can be argued that *all the points below the curve* (but above the abscissa) in Fig. 2.1 also belong to T. The premise behind this argument is the possibility of *free disposability of input* or – which is a reference to the same – that there are producers who are not as efficient regarding their production as the most efficient producers on the actual production function.

What about the points above the curve? Do any of these belong to T? No, if the data used in Table 2.1 derives from an *efficient producer* there will be no possibility – with the technology under consideration – of achieving yields above the curve in Fig. 2.1. However, if the data used in Table 2.1 derives from a “poor” producer – a producer who, if he had been a little more meticulous with his production, would have produced a higher yield at each of the indicated input levels – then there would have been points above the curve in Fig. 2.1 belonging to T, as T includes the points where x can produce y. And if this is a matter of only having received data from a “poor” producer, and a “good” producer would have been able to achieve a higher yield, then there would in fact be points above the curve in Fig. 2.1 belonging to T.

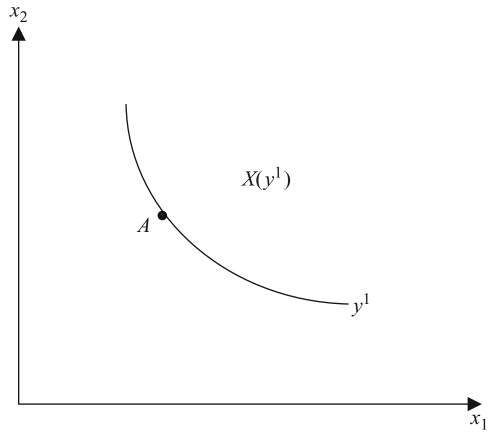
The problem is not insignificant and may give rise to considerable problems and challenges in connection with production economic research that makes use of empirical data (data from the real world). As it is, such data come from producers who are different, some of whom are “good” while others are “poor”. This being the case, the challenge is to establish which of these data do in fact make up the “border” of T (efficient producers) and which data derive from producers below the curve. The extent of the problem grows when the number of inputs (and outputs) increases to more than one.

This concludes the discussion of this issue. Notice that if the upper limit of T should be identical with the production function as illustrated in Fig. 2.1, then it presupposes that the data for the description of this function derives from an efficient producer.

2.3 The Input Set

The technology set T illustrates all the possible combinations of input and output. However, the production technology can also be defined in another way, i.e. as the *input set* $X(y)$ which is defined as:

Fig. 2.2 Isoquant and input set



$$X(y) = \{x : x\,\text{can produce}\,y\} \tag{2.2}$$

The input set $X(y)$ attaches to each value of y the amounts of input x that can produce y. If Fig. 2.1 is used again with the choice of a y value, e.g. $y = 70$, it is obvious that the input amount of 90 kg N can produce 70 units of cereal crops, i.e. $90 \in X(70)$. However, if 90 kg N can produce 70 units of cereal crops then a larger amount of N can also produce 70 units of cereal crops when the precondition of free disposability of input is applied. Hence, the set of x's which can produce 70 units of cereal crops consists of those amounts where $x \geq 90$, i.e. $X(70) = \{x: x \geq 90\}$. If all values of y are considered it would result in an illustration of the same technology sets as described in T.

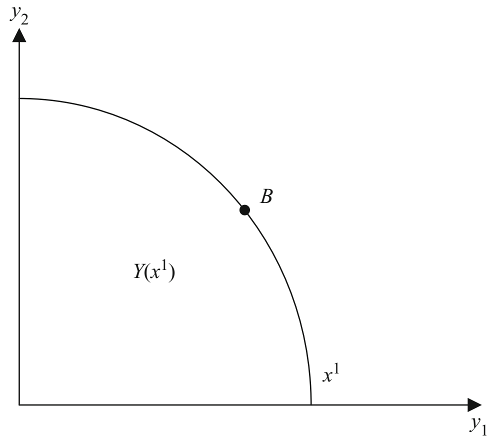
The input set can also be illustrated graphically when there are two inputs. Figure 2.2 shows an isoquant for production of the product y in the amount y^1 using the two inputs x_1 and x_2. An isoquant consists of those combinations of x_1 and x_2 that can produce the given product amount y^1. Hence, the point A illustrates an input combination which can in fact produce the amount y^1.

However, if such amounts of x_1 and x_2 – for instance corresponding to point A in Fig. 2.2 – can produce y^1 then larger amounts of x_1 and x_2 will also be able to produce the amount y^1 on the precondition of the existence of free disposability of input. Hence, the amounts that can produce y^1 $(X(y^1))$ are equal to the x's that are placed on and north-east of the isoquant in the figure.

2.4 The Output Set

The production technology can also be described by considering the product amounts which may be produced by a given input amount $x = x^1$. *The output set* $Y(x)$ is defined as:

Fig. 2.3 Production possibility curve and the output set



$$Y(x) = \{y : x \text{ can produce } y\} \tag{2.3}$$

The output set $Y(x)$ attaches to each value of x the amount of outputs that can be produced by use of the given amount of inputs. In Fig. 2.1, the amount of outputs that can be produced using 60 kg of nitrogen equals 45 units of cereal crops, i.e. $45 \in Y(60)$ in any case. However, if it is possible to produce 45 units of cereal crops with 60 kg of nitrogen, then it is also possible to produce *smaller* amounts of output with 60 kg of nitrogen. It would under all circumstances still be possible to produce the 45 units of cereal crops and then subsequently dispose of a part of the produced amount! Hence, on the precondition of *free disposability of output*, the set of y's that can be produced with 60 kg of nitrogen consists of those amounts where $y \leq 45$, i.e. $Y(60) = \{y: y \leq 45\}$.

The output set can also be illustrated graphically when there are two outputs. Figure 2.3 shows a production possibility curve for the production of the two products y_1 and y_2 with a given input amount x^1. The production possibility curve consists of those combinations of y_1 and y_2 that can be produced with a given input amount x^1. Hence, point B illustrates the output combination that can be produced with the input amount x^1.

However, if the amounts of y_1 and y_2 corresponding to point B can be produced by x^1, then smaller amounts of y_1 and y_2 could also be produced by the input amount x^1 on the precondition of the existence of free disposability of output. Hence, the amounts that can be produced by x^1 ($Y(x^1)$) are equal to the y's that are placed on and south-west of the production possibility curve and limited by the coordinate system axes.

2.5 The Production Function

With the definitions of the technology set, input set and output set presented in the above section in place, it is now possible to give a more formal and precise definition of a production function than the definition associated with the "empirical" production function described in Fig. 2.1. The following definition presupposes that y is a scalar (an output), while x is a scalar or a vector of input:

Definition. A production function f is defined as:

$$f(x) = \max\{y : y \in Y(x)\} \tag{2.4}$$

The production function could also be defined as:

$$f(x) = \max\{y : y \in \mathrm{T}(x, y)\} \tag{2.5}$$

Hence, a production function is defined as the maximum amount of output that can be produced (through the use of a given production technology) with a given amount of input.

Similarly, isoquants and production possibility curves can be given formal definitions. An *isoquant is defined* as "the border" of the input set, i.e. as the x's for which the following is true:

$$G(y) = \{x : x \in X(y) \mid x^k \notin X(y) \text{ for } x^k \leqslant x\} \tag{2.6}$$

in which $x^k \leq x$ is to be understood as: None of the elements (x_i) in the vector x^k are greater than the corresponding elements in the vector x, and at least one of the elements in x^k is smaller than the similar element in x.

If the possibility of production of multiple outputs exists, then the production possibility curve is defined similarly as:

$$P(x) = \{y : y \in Y(x) \mid y^k \notin Y(x) \text{ for } y^k \geqslant y\} \tag{2.7}$$

in which $y^k \geq y$ is to be understood as: None of the elements (y_i) in the vector y^k are smaller than the corresponding elements in the vector y, and at least one of the elements in y^k is greater than the similar element in y.

2.6 Diminishing Marginal Returns

Following this strictly formal definition of the production technology and production function, we shall now return to the graphical illustration of the production function which was the point of reference in Fig. 2.1 at the beginning of the chapter. But what would a purely graphical version of the production function look like? And what about the mathematical representation of the production function? What kinds of functions are used to represent production functions?

2.6.1 The Law of Diminishing Marginal Returns

First we will have a look at the graphical representation of a production function.

Recall that a production function can only be drawn on a piece of paper if there is one or at the most two inputs. As more than two inputs are normally used in a production, (almost) all graphical illustrations of production functions presuppose the presence of one or more underlying inputs (part of the production) with given fixed amounts (fixed input). The curve illustrating the relationship between added nitrogen fertiliser and the yield of cereal crops in Fig. 2.1, therefore, presupposes that all the other inputs used in the production of cereal crops (seeds, pesticides, land, labour, machinery, etc.) are present in given fixed amounts.

An essential precondition related to a production function is the assumption of *diminishing marginal returns*. The precondition, which is based on empirical observations of how the production is carried out in practice, is universally acknowledged as a basic condition within production economics referred to as the *Law of diminishing marginal returns*. Briefly explained,

> *The Law of diminishing marginal returns* states that by adding increasing amounts of input to a production with at least one fixed input, the additional returns resulting from the addition of increasing amounts of input will gradually diminish, and eventually become negative.

The concept of marginal returns is used here to refer to the increase in production arising from the addition of an extra unit of input. Normally, this increase is expressed by the slope of the production function, i.e. as the value of the derivative, i.e. $\mathrm{d}f(x)/\mathrm{d}x$, if x is a scalar, or the partial derivative, $\partial f(x)/\partial x_i$, if x is a vector. Expressed this way, the concept of *marginal returns* or *marginal product* is normally used to express the additional returns per input unit in connection with *marginal* changes in the amount of input.

If the function expression of the production function is unknown, the marginal product can be approximated by the use of the *difference product* expressed as $\Delta y/\Delta x$. Using data from the example in Table 2.1, the difference product in the interval from 30 to 60 kg of nitrogen equals (45–25)/(60–30) = 0.67, and in the interval from 90 to 120 kg of nitrogen equals (85–70)/(120–90) = 0.50. These difference products are approximated expressions of the derivative (and thereby the marginal product) at the centre of the relevant intervals.

The Law of diminishing marginal returns is nicely illustrated in the production function shown in Fig. 2.1. When adding small amounts of nitrogen fertiliser, the marginal product increases (the slope of the production function increases). At some point, the marginal product is diminishing, and when adding approximately 135 kg of nitrogen, the marginal product becomes zero and subsequently becomes negative with further additions. In this example, the precondition of at least one fixed input is satisfied as land and other inputs used in the production of cereal crops are presupposed to be present in given fixed amounts.

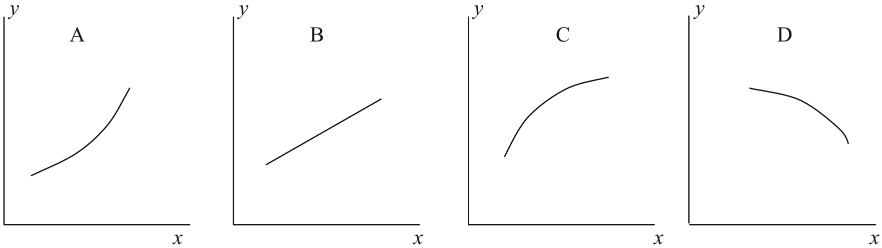


Fig. 2.4 Alternative production function shapes

2.6.2 Graphical Illustration of the Production Function

When production functions are represented graphically (and it is thereby presupposed that a number of underlying production factors are fixed inputs), such a representation will look the same as, or similar to, the curve in Fig. 2.1. These "similar" representations are produced by observing only parts of the shape of the total production function in Fig. 2.1.

Figure 2.4 illustrates four different (sub) shapes of the production function which are all contained in the production function outlined in Fig. 2.1. Example A outlines the progressively increasing shape with positive and increasing marginal returns. This shape corresponds to the first part of the production function in Fig. 2.1. Example B outlines a linear shape with positive and constant marginal returns, corresponding to the area between 60 and 90 kg N in Fig. 2.1. Example C outlines a digressively increasing shape with positive and diminishing marginal returns corresponding to the area between 90 and 130 kg N in Fig. 2.1. Finally, example D outlines a progressively diminishing shape with negative and diminishing marginal returns. This shape corresponds to the last part of the production function in Fig. 2.1.

A production function with all four "shape" types in the described order, like the one in Fig. 2.1, is referred to as the *neoclassical production function*. This type of production function has especially been used to describe production relationships within agriculture.

If you are not interested in the *overall shape* of the production function, but solely in the *local areas* of the production function, it is sufficient to plot the part of the production function that is of interest. As mentioned later on, the part of the production function that is of special interest in connection with production economics is the one that is illustrated in example C in Fig. 2.4 (the digressively increasing shape). Therefore, production functions are often illustrated graphically with a shape similar to example C in Fig. 2.4. However, this does not necessarily mean that this shape is present throughout the entire domain of the production function, i.e. *globally*. It might also solely be an issue of a description of a local shape.

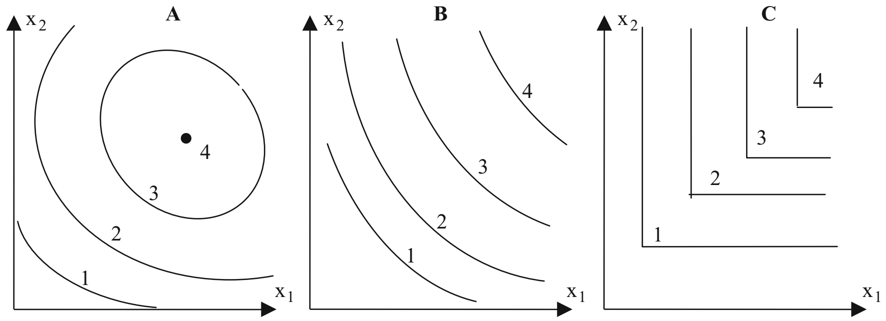


Fig. 2.5 Alternative sets of isoquants

If we consider production with more than one input, the graphical illustration of the production function is a bit more complicated. If there are two variable inputs, the production function is often described by means of so-called *isoquants* which are defined formally (for any number of inputs) in Eq. 2.6 and illustrated graphically as in Fig. 2.2 for two inputs. Isoquants can be interpreted as level curves for the production function. As the issue of interest regarding production economics is normally solely the area of the production function which corresponds to example C in Fig. 2.4, the similar areas of the isoquant will in fact consist of diminishing, convex curves, as illustrated in Fig. 2.2 (it is up to the reader to demonstrate why).

Figure 2.5 shows alternative sets of four isoquants. The number on each of the isoquants expresses the size of the production. In set A, the amount 1 can be produced with either input x_1 or with input x_2 or with a combination of x_1 and x_2. Hence, none of the inputs are *necessary inputs*. To produce amounts of 2, 3, or 4, both inputs are however necessary. The production function has a maximum of 4. In set B, both inputs are necessary and the production function does not have a maximum (this could e.g. be a Cobb-Douglas production function (discussed later)). Set C shows L-shaped isoquants on which only the corner points are efficient. This is a so-called Leontief production function (discussed later).

2.6.3 Mathematical Representation of the Production Function

The formal mathematical representation of the production function for the production of *one* output has previously been shown as in Eq. 2.4. Alternatively, Eq. 2.4 could be written as:

$$y = f(x) \tag{2.8}$$

in which y is a scalar (the amount of the product y), f is the production function, and x is a vector of inputs.

The production function:

$$y = f(x_1) \tag{2.9}$$

expresses the production of y only as a function of the variable input x_1. If it is appropriate to explicitly express that the production of output y is a function of the variable input x_1 and the fixed inputs $x_2, \ldots, x_n$, then the function (2.9) should be written as $y = f(x_1 \mid x_2, \ldots, x_n)$. Normally, fixed inputs are not included when writing the production function. It is however important to keep in mind that the production may depend on considerably more inputs than specified in the actual production function. Write $y = f(x_1, x_2)$ or $y = f(x_1, x_2, \mid x_3, \ldots, x_n)$ if you want to express that the production is a function of two variable inputs.

There is no given mathematical functional form for a production function. All the functional forms that have been used to describe the production have historically been based on more or less subjective choices. The best known of these function forms is the so-called Cobb-Douglas production function which, with two variable inputs, has the form:

$$y = Ax_1^{b_1}x_2^{b_2} \tag{2.10}$$

in which A, b_1, and b_2 are predetermined parameters (constants).

Evidently, the choice of functional form depends on the areas of the production function which are to be described. Is it a global description which should cover the entire function shape as outlined in Fig. 2.1, or is it a matter of functions which should only illustrate local areas of the production shape, as e.g. illustrated by the four examples in Fig. 2.4? Hence, the Cobb-Douglas function is only capable of illustrating shapes such as the one shown in example C in Fig. 2.4. An alternative functional form, which also seems to be able to work here, is the simple quadratic function. In case of a linear shape as shown in example B in Fig. 2.4, it is possible to choose a simple linear function as the functional form.

The choice of functional form and the subsequent estimation of the parameters of the function is a comprehensive science in itself, which is not discussed in any further detail here. Anyone with a particular interest in this is referred to studies within the subject area of Econometrics.

2.6.4 The Production Elasticity

Apart from describing the production technology as a table with numerical relationships between inputs and outputs (Table 2.1), as a graph illustrating these numerical relationships (Fig. 2.1), or mathematically as an actual production function (Eqs. 2.9 and 2.10), it is possible to express these relationships between inputs and outputs locally by means of the so-called *production elasticity*.

The production elasticity expresses the relative change in production through a relative change in the addition of input. If e.g. 5 % more input is added and 4 % more output is achieved, then the production elasticity is 4/5 = 0.80.

If there are multiple inputs, it is possible to calculate the production elasticity for each input. Formally, the production elasticity ε_i for input i is calculated as:

$$\varepsilon_i \equiv \frac{\frac{\partial f(\mathbf{x})}{f(\mathbf{x})}}{\frac{\partial x_i}{x_i}} = \frac{\frac{\partial f(\mathbf{x})}{\partial x_i}}{\frac{f(\mathbf{x})}{x_i}} = \frac{MPP_i}{APP_i} \tag{2.11}$$

in which MPP and APP represent the marginal product (*M*arginal *P*hysical *P*roduct) and the average product (*A*verage *P*hysical *P*roduct), respectively.

If the function expression for the production function is not known, then the production elasticity can be approximated by replacing marginal change (∂) in Eq. 2.11 by small, numerical change (Δ). Hence, an approximated expression for the production elasticity in the centre of the interval is achieved by calculating the following:

$$\varepsilon_i \cong \frac{\frac{\Delta y}{y}}{\frac{\Delta x_i}{x_i}}$$

If the data from the example in Table 2.1 is used, the production elasticity in the interval 30–60 kg N is approximated using the calculation ε_i = [(45–25)/25]/[(60–30)/30] = 0.80. As the centre of the interval 30–60 kg is 45 kg, this elasticity (0.80) will be used as the approximated elasticity at the point where the 45 kg N are applied. Similarly, the elasticity at the point where the 105 kg N are applied is expressed as ε_i = [(85–70)/70]/[(120–90)/90] = 0.64. As shown by this example, the production elasticity (normally) depends on the point of reference, and the elasticity declines with the addition of input.

Some production functions have constant production elasticities. This is the case for the Cobb-Douglas production function shown in Eq. 2.10 in which the production elasticity for the input i (i = 1,2) is b_i (the reader is encouraged to verify this himself/herself using the expression after the second equal sign in Eq. 2.11 for the calculation).

http://www.springer.com/978-3-642-30199-5

Production Economics
The Basic Theory of Production Optimisation
Rasmussen, S.
2013, XII, 292 p., Hardcover
ISBN: 978-3-642-30199-5