What is isoquant and isocost line

A rational firm seeks maximization of its profit. Maximization of profit implies minimization of cost. The cost is minimum, when the input combination is optimal. Therefore, choosing the right input combination leads to cost minimization and hence ensures maximum profits. In the theory of consumer behavior, we analyze the equilibrium of a consumer with the help of the indifference curve analysis. Similarly, the producer's equilibrium, which represents the least cost combination of inputs, can be examined with the help of isoquants. It should be remembered that the isoquants in the theory of production are, in fact, the counterpart of the indifference curves in the theory of consumption.AssumptionsThe principle of least-cost combination rests on the following assumptions:Capital and labor are the two factors involved in production. All the units of both the factors are homogeneous The prices of the factor units are given The total money outlay is also given There is perfect competition in the factor-market. In order to analyze producer's equilibrium, the firm under consideration should know its isoquant map an its isocost line.Isoquant MapIsoquants indicate various possibilities of combining two inputs. For each level of output, there will be a different isoquant. When a set of isoquants are depicted on a graph it is called an isoquant map.Isocost LineThe concept of isocost line is not a new one. It is the counterpart of the budget line in the theory of consumption. The isocost line is the producers' resource line.In the words of Prof. Barthwal, "it is a locus of all combinations of two (or more) inputs which the producer can buy using his fixed outlay at fixed input prices."We shall now draw the isocost line on the basis of an imaginary example.Let us assume that a firm has a sum of $500 to spend on two factors, labor and capital. Further, let us assume that a unit of labor costs $10 and a unit of capital (machine) costs $100. With the total outlay of $500, the firm could hire 50 units of labor and no capital, or it could hire 5 units of capital and no labor; or some combination of labor and capital in between. OM in the diagram represents 50 workers and ON represent 5 machines.If we connect the two points N and M, we get the isocost line. Thus, an isocost line gives all combinations of labor and capital at equal cost. The isocost line will shift when the prices of factors change, the outlay remaining the same. Likewise, the isocost line will shift to the right if the outlay of the firm increases. Hiring more of both inputs will cost more. When the total outlay is $600, the isocost line is N1M1. N2M2 is the isocost line when the outlay is $700.Thus, the isocost line depends upon two things:1. The prices of the factors of production2. The total outlay which the firm wants to make on the two factors of production.The slope of an isocost line is PL/PK, which is the ratio of the price of labor to the price of capital, when labor is shown on the X-axis and capital is shown on the Y-axis.The Optimum Combination of InputsLet us consider the geometry of the producer's equilibrium.Now the problem confronting the firm is to reach the highest possible isoquant with its given isocost line. In other words, it is the problem of getting the highest amount of output from the given outlay. Towards this end, the equal product map has been super imposed on the isocost line NM.In figure 2, NM is the firm's isocost line. Isoquants IQ1, IQ2 and IQ3 represent different levels of output. Equilibrium is attained at the point where the isoquant is tangent to the isocost line. The isocost line NM sets the upper boundary for the purchase of the inputs when outlay and input prices are given.Outlay is not sufficient to move to IQ3. Likewise, the segments of isoquants falling below the isocost line indicate under-utilization of his outlay fully. Rationality on the part of the producer requires full utilization of resources for optimization of output.Points A and B also satisfy the tangency condition and they lie within the reach of the producer. However, at these points the firm remains at a lower isoquant IQ1, which yields a lesser level of output than that on IQ2. Thus, E is the point of equilibrium from where there is no tendency on the part of the producer to move away. The firm will get its maximum output when it employs OL0 units of labor and OK0 units of capital. The equilibrium position of the firm can also be explained in terms of the equality between MRTS and the factor price ratio. The slope of the isoquant is the marginal rate of technical substitution (MRTS) and the slope of the isocost line indicates the factor price ratio. It follows that while in equilibrium,MRTSLK = PL/PKThus, marginal rate of technical substitution can also be written as the ratio of the marginal product of labor to that of the marginal product of capital.MPL/MPK = PL/PK or MPL/PL = MPK/PKCommentsJeegar Bhatt on September 13, 2015:Thank you very much. Very helpful explanation. The isocost and isoquant line are used to determine the cost minimising level of output produced by a competitive firm. Isoquant curve The isoquant is defined as all the feasible combinations of inputs used to produce a given level of output. In the illustration above we have an commodity which is produced using capital and labour. The illustrations depicts an outcome where 10 units of the commodity is being produced. We can see from the illustration that to produce 10 units of output, we can either use: 11 Units of capital and 3 units of labour5 units of capital and 5 units of labour3 units of capital and 11 units of labour The idea that the isoquant captures is that there is diminishing returns to each input. Isocost line The isocost line shows all the different combinations of inputs that could be purchased for a given level of costs. The illustration above depicts an isocost curve where the total cost is $10, the wage rate is $1 and the rental rate of capital is $1. In this case, we can see that we can purchase: 10 units of capital and 0 units of labor5 units of capital and 5 units of labor0 units of capital and 10 units of labor An isoquant shows all combination of factors that produce a certain output An isocost show all combinations of factors that cost the same amount. Isocosts and isoquants can show the optimal combination of factors of production to produce the maximum output at minimum cost. An isoquant shows all the combination of two factors that produce a given output In this diagram, the isoquant shows all the combinations of labour and capital that can produce a total output (Total Physical Product TPP) of 4,000. In the above isoquant, this could be 20 capital and 18 labour or (more capital intensive) 9 capital and 35 labour. (more labour intensive An isoquant is usually shaped concave because of the law of diminishing returns. With fixed capital employing extra workers gives a declining increase in the marginal product (MP) Marginal rate of factor substitution The marginal rate of substitution is the amount of one factor (e.g. K) that can be replaced by one factor (e.g. L). If 2 units of capital could be replaced with one-factor labour, the MRS would be 2 Diminishing marginal rate of substitution If the firm employs 2 L and 40 K. Then employing one extra worker can enable it to save 10K. This is quite an efficient saving. The firm only has to pay one extra worker but can save the cost of 40. However, at a combination of 9 Labour, employing an extra worker enables a saving of only 2 capital. Therefore, the more that workers are employed, there is a diminishing rate at which you can substitute the other factor. There comes a point, where employing more workers barely saves any capital at all. This is when diminishing returns of labour is very high – workers effectively get in each other's way. As one moves down the isoquant, output remains the same. Therefore the output gained from employing more labour must equal the output lost from employing more capital. MPP (L) x ΔL = MPP (K) x ΔK This equation gives us Isoquant map An isoquant map shows different levels of output. For example I1 may show the combinations of capital and labour that can produce 4,000 TPP. I2 may show the combinations of capital and labour that can produce 5,000 TPP. I5 is a higher output than I4 In the short-term, a firm faces a trade-off along one particular isoquant. But, in the long-term, a firm can invest in increasing capital stock and produce at a higher output for the same quantity of labour. Isocost An isocost shows all the combination of factors that cost the same to employ. In this example, a unit of labour and capital cost £6,666 each. If we employ 30K and 30L, the total cost will be £200,000 + £200,000 If we employ 10 K and 50L, the total cost will be £66,666 +£333,333 = £400,000 Change in labour costs In this example, initially, the cost of labour and capital is both £5,000. (e.g. 60L = 60 x £5,000 = £300,000) However, if Labour cost rises to £10,000, then the isocost shifts to the left. Now, to keep cost at £300,000, a firm could only employ 30 workers (30 x £10,000) The slope of an isocost is therefore P_L / P_K Profit maximisation To maximise profits, a firm will wish to produce at the point of the highest possible isoquant and minimum possible isocost In this example, we have one isocost and three isoquants. With the isocost of £400,000 the maximum output a firm can manage would be a TPP of 4,000. If it produced at say 13 K and 48 Labour, it would only be able to produce a TPP of 3,500. A total TPP of 4,500 is currently not possible without increasing costs beyond £400,000 Profit maximisation – the least cost method of production Another way of seeking to maximise profits is to target an output of say 4,00 and then find the isocost with the lowest possible cost. In this case, the isocost which touches the tangential point of the TPP is a TC of £400,000. Related Profit Maximisation Types of Costs Slideshare uses cookies to improve functionality and performance, and to provide you with relevant advertising. If you continue browsing the site, you agree to the use of cookies on this website. See our User Agreement and Privacy Policy. Slideshare uses cookies to improve functionality and performance, and to provide you with relevant advertising. If you continue browsing the site, you agree to the use of cookies on this website. See our Privacy Policy and User Agreement for details. Read this article to learn about the laws of returns: the isoquant-isocost approach! The various production functions were explained in terms of the traditional analysis. This article explains them with the help of the isoquant-isocost approach. Image Courtesy : www2.econ.iastate.edu/classes/econ101/choi/img%5Chol004.jpg The technique involved here is similar to the indifference curve technique used in consumption theory. Isoquants: An isoquant (isoproduct) is a curve on which the various combinations of labour and capital show the same output. According to Cohen and Cyert, "An isoproduct curve is a curve along which the maximum achievable rate of production is constant." It is also known as a production indifference curve or a constant product curve. Just as indifference curve shows the various combinations of any two commodities that give the consumer the same amount of satisfaction (iso-utility), similarly an isoquant indicates the various combinations of two factors of production which give the producer the same level of output per unit of time. Table 24.1 shows a hypothetical isoquant schedule of a firm producing 100 units of a good. TABLE 24.1: Isoquant Schedule: Combination Units of Capital Units of Labour Total Output (in units) A 9 5 100 B 6 10 100 C 4 15 100 D 3 20 100 This Table 24.1 is illustrated on Figure 24.1 where labour units are measured along the X-axis and capital units on the K-axis. The first, second, third and the fourth combinations are shown as A, S, C and D respectively. Connect all these points and we have a curve IQ. This is an isoquant. The firm can produce 100 units of output at point A on this curve by having a combination of 9 units of capital and 5 units of labour. Similarly, point B shows a combination of 6 units of capital and 10 units of labour; point C,4 units of capital and' 15 units of labour; and point D, a combination of 3 units of capital and 20 units of labour to yield the same output of 100 units. An isoquant map shows a number of isoquants representing different amounts of output. In Figure 24.1, curves IQ, IQ1 and IQ2 show an isoquant map. Starting from the curve IQ which yields 100 units of product, the curve IQ1, shows 200 units and the IQ2 curve 300 units of the product which can be produced with altogether different combinations of the two factors. Isoquants vs. Indifference Curves: An isoquant is analogous to an indifference curve in more than one way. In it, two factors (capital and labour) replace two commodities of consumption. An isoquant shows equal level of product while an indifference curve shows equal level of satisfaction at all points. The properties of isoquants, as we shall study below, are exactly similar to those of indifference curves. However, there are certain differences between isoquants and indifference curves. Firstly, an indifference curve represents satisfaction which cannot be measured in physical units. In the case of an isoquant the product can be measured in physical units. Secondly, on an indifference map one can only say that a higher indifference curve gives more satisfaction than a lower one, but it cannot be said how much more or less satisfaction is being derived from one indifference curve as compared to the other, whereas one can easily tell by how much output is greater on a higher isoquant in comparison with a lower isoquant. In Figure 24.1 output on the curve 1QI is double, and on the IQ2 treble than on the curve IQ. Lastly, since satisfaction on indifference curves cannot be measured in physical units, they are given arbitrary numbers 1, 2, 3, 4, etc. The isoquants have an added advantage over the former because they can be labelled in physical units, as 100, 200, 300, etc. in Figure 24.1, to indicate the output level to which each curve corresponds. Properties of Isoquants: Isoquants possess certain properties which are similar to those of indifference curves. (1) Isoquants are negatively inclined: If they do not have a negative slope, certain logical absurdities follow. If the isoquant slopes upward to the right, it implies that both capital and labour increase but they produce the same output. In Figure 24.2 (A), combination B on the IQ curve having a larger amount of both capital and labour (OC1 +OL1 > OC + OL) will yield more output than before. Therefore, point A and B on the IQ curve cannot be of equal product. Suppose the isoquant is vertical as shown in Figure 24.2 (B), which implies a given amount of labour is combined with different units of capital. Since OL of labour and OC1 of capital will produce a larger amount than produced by OL of labour and OC of capital, the isoquant IQ cannot be a constant product curve. Take Figure 24.2 (C) where the isoquant is horizontal which means combining more of labour with the same quantity of capital. Here OC of capital and OL1 of labour will produce a larger or smaller amount than produced by the combination OC of capital and OL of labour. Therefore, a horizontal isoquant cannot be an equal product curve. Thus it is clear that an isoquant must slope downward to the right as shown in Figure 24.2 (D) where points A and B on the IQ curve are of equal quantity. As the amount of capital decreases from OC to OC1 and that of labour increases from OL to OL1 so that output remains constant. (2) An Isoquant lying above and to the right of another represents a higher output level. In Figure 24.3 combination B on IQ1 curve shows larger output than point A on the curve IQ. The combination of OC of capital and OL of labour yields 100 units of product while OC1 of capital and OL1 of labour produce 200 units. Therefore, the isoquant IQ1 which lies above and to the right of the isoquant IQ, represents a larger output level. (3) No two isoquants can intersect each other. The absurd conclusion that follows when two isoquants cut each other is explained with the aid of Figure 24.4. On the isoquant IQ, combination A =B. And on the isoquant IQ1 combination R=S. But combination S is preferred to combination B, being on the higher portion of isoquant IQ1. On the other hand, combination A is preferred to R, the former being on the higher portion of the isoquant IQ. To put it algebraically, it means that S> B and R< A. But this is logically absurd because S combination is as productive as R and A combination produces as much as B. Therefore, the same combination cannot both be less and more productive at the same time. Hence two isoquants cannot intersect each other. (4) Isoquants need not be parallel because the rate of substitution between two factors is not necessarily the same in all the isoquant schedules. (5) In between two isoquants there can be a number of isoquants showing various levels of output which the combinations of the two factors can yield. In fact, in between the units of output 100, 200, 300, etc. represented on isoquants there can be innumerable isoquants showing 120, 150, 175,235, or any other higher or lower unit. (6) Units of output shown on isoquants are arbitrary. The various units of output such as 100, 200, 300, etc., shown in an isoquant map are arbitrary. Any units of output such as 5, 10, 15, 20 or 1000, 2000, 3000, or any other units can be taken. (7) No isoquant can touch either axis. If an isoquant touches X-axis, it would mean that the product is being produced with the help of labour alone without using capital at all. This is a logical absurdity for OL units of labour alone are incapable of producing anything. Similarly, OC units of capital alone cannot produce anything without the use of labour. Therefore IQ and lQ1 cannot be isoquants, as shown in Figure 24.5. (8) Each isoquant is convex to the origin: As more units of labour are employed to produce 100 units of the product, lesser and lesser units of capital are used. This is because the marginal rate of substitution between two factors diminishes. In Figure 24.6, in order to produce 100 units of the product, as the producer moves along the isoquant from combination A to B and to C and D, he gives up smaller and smaller units of capital for additional units of labour. To maintain the same output of 100 units, BR less of capital and relatively RC more of labour is used. If he were producing this output with the combination D, he would be employing CT less of capital and relatively TD more of labour. Thus the isoquants are convex to the origin due to diminishing marginal rate of substitution. This fact becomes clear from successively smaller triangles below the IQ curve Δ ASB > ΔBRC > Δ CTD. (9) Each isoquant is oval-shaped: It is elliptical which means that at some point it begins to recede from each axis. This shape is a consequence Labour of fact that if a producer uses more of capital or more of labour or more Fig. 24.6 of both than is necessary, the total product will eventually decline. The firm will produce only in those segments of the isoquants which are convex to the origin and lie between the ridge lines. This is the economic region of production. In Figure 24.7, oval-shaped isoquants are shown. Curves OA and OB are the ridge lines and in between them economically feasible units of capital and labour can be employed to produce 100, 200, 300 and 400 units of the product. For example, OT units of labour and ST units of the capital can produce 100 units of the product, but the same output can be obtained by using the same quantity of labour OT and less quantity of capital VT. Thus only an unwise entrepreneur will produce in the dotted region of the isoquant 100. The dotted segments of an isoquant are the waste- bearing segments. They form the uneconomic regions of production. In the upper dotted portion, more capital and in the lower dotted portion more labour than necessary is employed. Hence GH, JK, LM, and NP segments of the elliptical curves are the iso- quants. Isocost Curves: Having studied the nature of isoquants which represent the output possibilities of a firm from a given combination of two inputs, we pass on to the prices of the inputs as represented on the isoquant map by the isocost curves. These curves are also known as outlay lines, price lines, input-price lines, factor-cost lines, constant-outlay lines, etc. Each isocost curve represents the different combinations of two inputs that a firm can buy for a given sum of money at the given price of each input. Figure, 24.8 (A) shows three isocost curves AB, CD and EF, each represents a total outlay of 50, 75 and 100 respectively. The firm can hire OC of capital or OD of labour with Rs. 75. OC is 2/3 of OD which means that the price of a unit of labour is 1½ times less than that of a unit of capital. The line CD represents the price ratio of capital and labour. Prices of factors remaining the same, if the total outlay is raised, the isocost curve will shift upward to the right as EF parallel to CD, and if the total outlay is reduced it will shift downwards to the left as AB. The isocosts are straight lines because factor prices remain the same whatever the outlay of the firm on the two factors. The isocost curves represent the locus of all combinations of the two input factors which result in the same total cost. If the unit cost of labour (L) is w and the unit cost of capital (C) is r, then the total cost: TC = wL + rC. The slope of the isocost line is the ratio of prices of labour and capital i.e., w/r. The point where the isocost line is tangent to an isoquant represents the least cost combination of the two factors for producing a given output. If all points of tangency like LMN are joined by a line, it is known as an output- factor curve or least-outlay curve or the expansion path of a firm. Salvatore defines expansion path as "the locus of points of producer's equilibrium resulting from changes in total outlays while keeping factor prices constant." It shows how the proportions of the two factors used might be changed as the firm expands. For example, in Figure 24.8 (A) the proportions of capital and labour used to produce 200 (IQ1) units of the product are different from the proportions of these factors used to produce 300 (IQ2) units or 100 (OQ) units at the lowest cost. Like the price-income line in the indifference curve analysis, a relative cheapening of one of the factors to that of another will extend the isocost line to the right. If one of the factors becomes relatively dearer, the isocost line will contract inward to the left. Given the price of capital, if the price of labour falls, the isocost line EF in Panel (B) will extend to the right as EG and if the price of labour rises, the isocost line EF will contract inward to the left as EH. if the equilibrium points L, M, and N are joined by a line, it is called the price-factor curve. The Principle of Marginal Rate of Technical Substitution: The principle of marginal rate of technical substitution (MRTS or MRS) is based on the production function where two factors can be substituted in variable proportions in such a way as to produce a constant level of output. The marginal rate of technical substitution between two factors C (capital) and L (labour), MRTSLC is the rate at which L can be substituted for C in the production of good X without changing the quantity of output. As we move along an isoquant downward to the right, each point on it represents the substitution of labour for capital. MRTS is the loss of certain units of capital which will just be compensated for by additional units of labour at that point. In other words, the marginal rate of technical substitution of labour for capital is the slope or gradient of the isoquant at a point. Accordingly, slope = MRTSLC = – Δ C/A L. This can be understood with the aid of the isoquant schedule, in Table 24.2. TABLE 24.2: Isoquant Schedule: Combination Labour Capital MRTSLC Output 1 5 9 __ 100 2 10 6 3:5 100 3 15 4 2:5 100 4 20 3 L;5 100 The above table shows that in the second combination to keep output constant at 100 units, the reduction of 3 units of capital requires the addition of 5 units of labour, MRTSLC = 3:5. In the third combination, the loss of 2 units of capital is compensated for by 5 more units of labour, and so on. In Figure 24.9 at point B, the marginal rate of technical substitution is AS/SB, at point G, it is BT/TG and at H, it is GR/ RH. The isoquant AH reveals that as the units of labour are successively increased into the factor-combination to produce 100 units of good X, the reduction in the units of capital becomes smaller and smaller. It means that the marginal rate of technical substitution is diminishing. This concept of the diminishing marginal rate of technical substitution (DMRTS) is parallel to the principle of diminishing marginal rate of substitution in the indifference curve technique. This tendency of diminishing marginal substitutability of factors is apparent from Table 24.2 and Figure 24.9. The MRTSLC continues to decline from 3:5 to 1:5 whereas in the Figure 24.9 the vertical lines below the triangles on the isoquant become smaller and smaller as we move downward so that GR < BT < AS. Thus, the marginal rate of technical substitution diminishes as labour is substituted for capital. It means that the isoquant must be convex to the origin at every point. The Law of Variable Proportions: The behaviour of the law of variable proportions or of the short-run production function when one factor is constant and the other variable can also be explained in terms of the isoquant analysis. Suppose capital is a fixed factor and labour is a variable factor. In Figure 24.10., OA and OB are the ridge lines and it is in between them that economically feasible units of labour and capital can be employed to produce 100, 200, 300, 400 and 500 units of output. It implies that in these portions of the isoquants, the marginal product of labour and capital is positive. On the other hand, where these ridge lines cut the isoquants, the marginal product of the inputs is zero. For instance, at point H the marginal product of capital is zero, and at point L the marginal product of labour is zero. The portion of the isoquant that lies outside the ridge lines, the marginal product of that factor is negative. For instance, the marginal product of capital is negative at G and that of labour at R. The law of variable proportions says that, given the technique of production, the application of mote and more units of a variable factor, say labour, to a fixed factor, say capital, will, until a certain point is reached, yield more than proportional increases in output, and thereafter less than proportional increases in output. Since the law refers to increases in output, it relates to the marginal product. To explain the law, capital is taken as a fixed factor and labour as a variable factor. The isoquants show different levels of output in the figure. OC is the fixed quantity of capital which therefore forms a horizontal line CD. As we move from C to D towards the right on this line, the different points show the effects of the combinations of successively increasing quantities of labour with fixed quantity of capital OC. To begin with, as we move from C to G to H, it shows the first stage of increasing marginal returns of the law of variable proportions. When CG labour is employed with OC capital, output is 100. To produce 200 units of output, labour is increased by GH while the amount of capital is fixed at OC. The output has doubled but the amount of labour employed has not increased proportionately. It may be observed that GH < CG, which means that smaller additions to the labour force have led to equal increment in output. Thus C to H is the first stage of the law of variable proportions in which the marginal product increases because output per unit of labour increases as more output is produced. The second stage of the law of variable proportions is the portion of the isoquants which lies in between the two ridge lines O A and OB. It is the stage of diminishing marginal returns between points H and L. As more labour is employed, output increases less than proportionately to the increase in the labour employed. To raise output to 300 units from 200 units, HJ labour is employed. Further, JK quantity of labour is required to raise output from 300 to 400 and KL of labour to raise output from 400 to 500. So, to increase output by 100 units successively, more and more units of the variable factor (labour) are required to be applied along with the fixed factor (capital) , that is KL>JK>HJ. It implies that the marginal product of labour continues to decline with the employment of larger quantities to it. Thus as we more from point H to K, the effect of increasing the units of labour is that output per unit of labour diminishes as more output is produced. This is known as the stage of diminishing returns. If labour is employed further, we are outside the lower ridge line OB and enter the third stage of the law of variable proportions. In this region which lies beyond the ridge line OB there is too much of the variable factor (labour) in relation to the fixed factor (capital). Labour is thus being overworked and its marginal product is negative. In other words when the quantity of labour is increased by LR and RS, the output declines from 500 to 400 and to 300. This is the stage of negative marginal returns. We arrive at the conclusion that a firm will find it profitable to produce only in the second stage of the law of variable proportions for it will be uneconomical to produce in the regions to the left or right of the ridge lines which form the first stage and the third stage of the law respectively. The Laws of Returns to Scale: The laws of returns to scale can also be explained in terms of the isoquant approach. The laws of returns to scale refer to the effects of a change in the scale of factors (inputs) upon output in the long-run when the combinations of factors are changed in some proportion. If by increasing two factors, say labour and capital, in the same proportion, output increases in exactly the same proportion, there are constant returns to scale. If in order to secure equal increases in output, both factors are increased in larger proportionate units, there are decreasing returns to scale. If in order to get equal increases in output, both factors are increased in smaller proportionate units, there are increasing returns to scale. The returns to scale can be shown diagrammatically on an expansion path "by the distance between successive 'multiple-level-of-output' isoquants, that is, isoquants that show levels of output which are multiples of some base level of output, e.g., 100, 200, 300, etc." Increasing Returns to Scale: Figure 24.11 shows the case of increasing returns to scale where to get equal increases in output, lesser proportionate increases in both factors, labour and capital, are required. It follows that in the figure: 100 units of output require 3C +3L 200 units of output require 5C + 5L 300 units of output require 6C + 6L So that along the expansion path OR, OA > AB > BC. In this case, the production function is homogeneous of degree greater than one. The increasing returns to scale are attributed to the existence of indivisibilities in machines, management, labour, finance, etc. Some items of equipment or some activities have a minimum size and cannot be divided into smaller units. When a business unit expands, the returns to scale increase because the indivisible factors are employed to their full capacity. Increasing returns to scale also result from specialisation and division of labour. When the scale of the firm expands there is wide scope for specialisation and division of labour. Work can be divided into small tasks and workers can be concentrated to narrower range of processes. For this, specialized equipment can be installed. Thus with specialization, efficiency increases and increasing returns to scale follow. Further, as the firm expands, it enjoys internal economies of production. It may be able to install better machines, sell its products more easily, borrow money cheaply, procure the services of more efficient manager and workers, etc. All these economies help in increasing the returns to scale more than proportionately. Not only this, a firm also enjoys increasing returns to scale due to external economies. When the industry itself expands to meet the increased 'long-run demand for its product, external economies appear which are shared by all the firms in the industry. When a large number of firms are concentrated at one place, skilled labour, credit and transport facilities are easily available. Subsidiary industries crop up to help the main industry. Trade journals, research and training centres appear which help in increasing the productive efficiency of the firms. Thus these external economies are also the cause of increasing returns to scale. Decreasing Returns to Scale: Figure 24.12 shows the case of decreasing returns where to get equal increases in output, larger proportionate increases in both labour and capital are required. It follows that: 100 units of output require 2C + 2L 200 units of output require 5C + 5L 300 units of output require 9C + 9L So that along the expansion path OR, OG < GH < HK. In this case, the production function is homogeneous of degree less than one. Returns to scale may start diminishing due to the following factors. Indivisible factors may become inefficient and less productive. The firm experiences internal diseconomies. Business may become unwieldy and produce problems of supervision and coordination. Large management creates difficulties of control and rigidities. To these internal diseconomies are added external diseconomies of scale. These arise from higher factor prices or from diminishing productivities of the factors. As the industry continues to expand the demand for skilled labour, land, capital, etc. rises. There being perfect competition, intensive bidding raises wages, rent and interest. Prices of raw materials also go up. Transport and marketing difficulties emerge. All these factors tend to raise costs and the expansion of the firms leads to diminishing returns to scale so that doubling the scale would not lead to doubling the output. Constant Returns to Scale: Figure 24.13 shows the case of constant returns to scale. Where the distance between the isoquants 100, 200 and 300 along the expansion path OR is the same, i.e., OD = DE = EF. It means that if units of both factors, labour and capital, are doubled, the output is doubled. To treble output, units of both factors are trebled. It follows that: 100 units of output require 1 (2C + 2L) = 2C + 2L 200 units of output require 2(2C + 2L) = 4C + 4L 300 units of output require 3(2C + 2L) = 6C + 6L The returns to scale are constant when internal economies enjoyed by a firm are neutralised by internal diseconomies so that output increases in the same proportion. Another reason is the balancing of external economies and external diseconomies. Constant returns to scale also result when factors of production are perfectly divisible, substitutable, homogeneous and their supplies are perfectly elastic at given prices. That is why, in the case of constant returns to scale, the production function is homogeneous of degree one. Relation between Returns to Scale and Returns to a Factor (Law of Returns to Scale and Law of Diminishing Returns): Returns to a factor and returns to scale are two important laws of production. Both laws explain the relation between inputs and output. Both laws have three stages of increasing, decreasing and constant returns. Even then, there are fundamental differences between the two laws. Returns to a factor relate to the short period production function when one factor is varied keeping the other factor fixed in order to have more output, the marginal returns of the variable factor diminish. On the other hand, returns to scale relate to the long period production function when a firm changes its scale of production by changing one or more of its factors. We discuss the relation between the returns to a factor (law of diminishing returns) and returns to scale (law of returns to scale) on the assumptions that: (1) There are only two factors of production, labour and capital. (2) Labour is the variable factor and capital is the fixed factor. (3) Both factors are variable in returns to scale. (4) The production function is homogeneous. Given these assumptions, we first explain the relation between constant return to scale and returns to a variable factor in terms of Figure 24.14 where OS is the expansion path which shows constant returns to scale because the difference between the two isoquants 100 and 200 on the expansion path is equal i.e., OM = MN. To produce 100 units, the firm uses OC + OL quantities of capital and labour and to double the output to 200 units, double the quantities of labour and capital are required so that OC1 + OL2 lead to this output level at point N. Thus there are constant returns to scale because OM = MN. To prove that returns to the variable factor, labour, diminish, we take OC of capital as the fixed factor, represented by the CC, line. Keeping C as constant, if the amount of labour is doubled by LL2, we reach point K which lies on a lower isoquant 150 than the isoquant 200. By keeping C constant, Ц if the output is to be doubled from 100 to 200 units, then L3 units o of labour will be required. But L3 > L2. Thus by doubling the units of labour with constant C2, the output less than doubles. It is 150 units at point K instead of 200 units at point P. This shows that the marginal returns of the variable factor, labour, have diminished. As pointed out by Stonier and Hague, "So, if production function were always homogeneous of the first degree and if returns to scale were always constant, marginal physical productivity (returns) would always fall." The relation between diminishing returns to scale and return to a variable factor is explained with the help of Figure 24.15 where OS is the expansion path which depicts diminishing returns to scale because the segment MN>OM. It means that in order to double the output from 100 to 200, more than double the amounts of both factors are required. Alternatively, if both factors are doubled to OC2+ OL2 they lead to the lower output level isoquant 175 at point R than the isoquant 200 which shows diminishing returns to scale. If C is kept constant and the amount of variable factor, labour, is doubled by LL2 we reach point K which lies on a still lower level of output represented by the isoquant 140. This proves that the marginal returns (or physical productivity) of the variable factor, labour, have diminished. 3. Now we take the relation between increasing returns to scale and returns to a variable factor. This is explained in terms of Figure 24.16 (A) and (B). In Panel (A), the expansion path OS depicts increasing returns to scale because the segment OM > MN. It means that in order to double the output from 100 to 200, less than double the amounts of both factors will be required. If C is kept constant and the amount of variable factor, labour, is doubled by LL2 the level of output is reached at point K which shows diminishing marginal returns as represented by the lower isoquant 160 than the isoquant 200 when returns to scale are increasing. In case the returns to scale are increasing strongly, that is, they are highly positive they will offset the diminishing marginal returns of the variable factor, labour. Such a situation leads to increasing marginal returns. This is explained in Panel (B) of Figure 24.16 where on the expansion path OS, the segment OM > MN, thereby showing increasing returns to scale. When the amount of the variable factor, labour, is doubled by LL2 while keeping C as constant, we reach the output level K represented by the isoquant 250 which is at a higher level than the isoquant 200. This shows that the marginal returns of the variable factor, labour, have increased even when there are increasing returns to scale. Conclusion: It can be concluded from the above analysis that under a homogeneous production function when a fixed factor is combined with a variable factor, the marginal returns of the variable factor diminish when there are constant, diminishing and increasing returns to scale. However, if there are strong increasing returns to scale, the marginal returns of the variable factor increase instead of diminishing. Choice of Optimal Factor Combination or Least Cost Combination of Factors or Producer's Equilibrium: A profit maximisation firm faces two choices of optimal combination of factors (inputs): First, to minimise its cost for a given output; and second, to maximise its output for a given cost. Thus the least cost combination of factors refers to a firm producing the largest volume of output from a given cost and producing a given level of output with the minimum cost when the factors are combined in an optimum manner. We study these cases separately. Cost-Minimisation for a Given Output: In the theory of production, the profit maximisation firm is in equilibrium when, given the cost-price function, it maximises its profits on the basis of the least cost combination of factors. For this, it will choose that combination which minimises its cost of production for a given output. This will be the optimal combination for it. Assumptions: This analysis is based on the following assumptions: 1. There are two factors, labour and capital. 2. All units of labour and capital are homogeneous. 3. The prices of units of labour (w) and that of capital (r) are given and constant. 4. The cost outlay is given. 5. The firm produces a single product. 6. The price of the product is given and constant. 7. The firm aims at profit maximisation. 8. There is perfect competition in the factor market. Given these assumptions, the point of least-cost combination of factors for a given level of output is where the isoquant curve is tangent to an isocost line. In Figure 24.17, the isocost line GH is tangent to the isoquant 200 at point M. The firm employs the combination of OC of capital and OL of labour to produce 200 units of output at point M with the given cost- outlay GH. At this point, the firm is minimising its cost for producing 200 units. Any other combination on the isoquant 200, such as R or T, is on the higher isocost line KP which shows higher cost of production. The isocost line EF shows lower cost but output 200 cannot be attained with it. Therefore, the firm will choose the minimum cost point M which is the least-cost factor combination for producing 200 units of output. M is thus the optimal combination for the firm. The point of tangency between the isocost line and the isoquant is an important first order condition but not a necessary condition for the producer's equilibrium. There are two essential or second order conditions for the equilibrium of the firm. 1. The first condition is that the slope of the isocost line must equal the slope of the isoquant curve. The Slope of the isocost line is equal to the ratio of the price of labour (w) to the price of capital (r) i.e., w/r. The slope of the isoquant curve is equal to the marginal rate of technical substitution of labour and capital (MRTSLC) which is, in turn, equal to the ratio of the marginal product of labour to the marginal product of capital (MPL/MPC). Thus the equilibrium condition for optimality can be written as: The second condition is that at the point of tangency, the isoquant curve must he convex to the origin. In other words, the marginal rate of technical substitution of labour for capital (MRTSLC) must be diminishing at the point of tangency for equilibrium to be stable. In Figure 24.18, S cannot be the point of equilibrium, for the isoquant IQ1, is concave where it is tangent to the isocost line GH. At point S, the marginal rate of technical substitution between the two factors increases if move to the right m or left on the curve lQ1. Moreover, the same output level can be produced at a lower cost CD or EF and there will be a corner solution either at C or F. If it decides to produce at EF cost, it can produce the entire output with only OF labour. If, on the other hand, it decides to produce at a still lower cost CD, the entire output can be produced with only OC capital. Both the situations are impossibilities because nothing can be produced either with only labour or only capital. Therefore, the firm can produce the same level of output at point M where the isoquant curve IQ is convex to the origin and is tangent to the isocost line GH. The analysis assumes that both the isoquants represent equal level of output, IQ = IQ1. Output-Maximisation for a Given Cost: The firm also maximises its profits by maximising its output, given its cost outlay and the prices of the two factors. This analysis is based on the same assumptions, as given above. The conditions for the equilibrium of the firm are the same, as discussed above. 1. The firm is in equilibrium at point P where the isoquant curve 200 is tangent to the isocost line CL. At this point, the firm is maximising its output level of 200 units by employing the optimal combination of OM of capital and ON of labour, given its cost outlay CL. But it cannot be at points E or F on the isocost line CL, since both points give a smaller quantity of output, being on the isoquant 100, than on the isoquant 200. The firm can reach the optimal factor combination level of maximum output by moving along the isocost line CL from either point E or F to point P. This movement involves no extra cost because the firm remains on the same isocost line. The firm cannot attain a higher level of output such as isoquant 300 because of the cost constraint. Thus the equilibrium point has to be P with optimal factor combination OM + ON. At point P, the slope of the isoquant curve 200 is equal to the slope of the isocost line CL. It implies that w/r=MPL/MPC=MRTSLC 2. The second condition is that the isoquant curve must be convex to the origin at the point of tangency with the isocost line, as explained above in terms of Figure 24.18.