PARTIAL DERIVATIVES

MATH 195, SECTION 59 (VIPUL NAIK)

Corresponding material in the book: Section 14.3.

What students should definitely get: Definition and computation techniques for first and higher partials, both at a specific point and as general expressions, statement of Clairaut's theorem.

What students should hopefully get: Cases of particular interest in partial derivative computation, interpretation of signs of partial derivatives.

Executive summary

Words ...

(1) The partial derivative of a function of many variables with respect to any one variable is the derivative with respect to that variable, keeping others constant. It can be written as a limit of a difference quotient, using variable letter subscript (such as $f_x(x, y)$), numerical subscript based on input position (such as $f_2(x_1, x_2, x_3)$), Leibniz notation (such as $\partial/\partial x$).
(2) In the separate continuity-joint continuity paradigm, partial derivatives correspond to the "separate" side. The corresponding "joint" side notion requires linear algebra and we will therefore defer it.
(3) The expression for the partial derivative of a function of many variables with respect to any one of them involves all the variables, not just the one being differentiated against (the exception is additively separable functions). In particular, the *value* of the partial derivative (as a number) depends on the values of all the inputs.
(4) The procedure for partial derivatives differs from the procedure used for implicit differentiation: in partial derivatives, we assume that the other variable is independent and constant, while in implicit differentiation, we treat the other variable as an unknown (implicit) function of the variable.
(5) We can combine partial derivatives and implicit differentiation, for instance, $G(x, y, z) = 0$ may be a description of z as an implicit function of x and y, and we can compute $\partial z/\partial x$ by implicit differentiation, differentiate G, treat z as an implicit function of x and treat y as a constant.
(6) By iterating partial differentiation, we can define higher order partial derivatives. For instance f_{xx} is the derivative of f_x with respect to x. For a function of two variables x and y, we have four second order partials: f_{xx}, f_{yy}, f_{xy} and f_{yx}.
(7) Clairaut's theorem states that if f is defined in an open disk surrounding a point, and both mixed partials f_{xy} and f_{yx} are jointly continuous in the open disk, then $f_{xy} = f_{yx}$ at the point.
(8) We can take higher order partial derivatives. By iterated application of Clairaut's theorem, we can conclude that under suitable continuity assumptions, the mixed partials having the same number of differentiations with respect to each variable are equal in value.
(9) We can consider a partial differential equation for functions of many variables. This is an equation involving the function and its partial derivatives (first or higher order) all at one point. A solution is a function of many variables that, when plugged in, satisfies the partial differential equation.
(10) Unlike the case of ordinary differential equations, the solution spaces to partial differential equations are huge, usually infinite-dimensional, and there is often no neat description of the general solution.

Pictures ...

(1) The partial derivatives can be interpreted as slopes of tangent lines to graphs of functions of the one variable being differentiated with respect to, once we fix the value of the other variable.

Actions ...

(1) To compute the first partials, differentiate with respect to the relevant variable, treating other variables as constants.

(2) Implicit differentiation for first partial of implicit function of two variables, e.g., z as a function of x and y given via $G(x, y, z) = 0$.
(3) In cases where differentiation formulas do not apply directly, use the limit of difference quotient idea.
(4) To calculate partial derivative at a point, it may be helpful to first fix the values of the other coordinates and then differentiate the function of one variable rather than trying to compute the general expression for the derivative using partial differentiation and then plugging in values. On the other hand, it might not.
(5) Two cases of particular note for computing partial derivatives are the cases of additively and multiplicatively separable functions.
(6) To find whether a function satisfies a partial differential equation, plug it in and check. Don't try to find a general solution to the partial differential equation.

Econ-speak ...

(1) Partial derivatives = marginal analysis. Positive = increasing, negative = decreasing
(2) Second partial derivatives = nature of returns to scale. Positive = increasing returns (concave up), zero = constant returns (linear), negative = decreasing returns (concave down)
(3) Mixed partial derivatives = interaction analysis; positive mixed partial derivative means complementary, negative mixed partial derivative means substitution
(4) The signs of the first partials are invariant under monotone transformations, not true for signs of second partials, pure or mixed.
(5) Examples of quantity demanded, production functions.
(6) Cobb-Douglas production functions (see section of lecture notes and corresponding discussion in the book)

1. Partial derivatives: introduction

If there's one concept that is really unique to multivariable calculus, and really important at a conceptual level in applications of mathematics to the social sciences, it is the concept of partial derivatives. This is particularly true in the case of economics, since a key ingredient of economic thinking is *marginal analysis*, which is just another way of saying *partial derivatives*. The truly revolutionary idea of *mixed partial derivatives* is fairly important, and serves as a conceptual lens for understanding the interaction of multiple variables.

1.1. Separate versus joint, and partial derivatives.

Recall, from our discussion of limits and continuity for functions of many variables, the concept of *separate* versus *joint*. Separate continuity referes to continuity in each of the variables in isolation, where we are allowed to move only one variable at a time. Joint continuity refers to continuity where we are allowed to simultaneously move more than one variable, i.e., it refers to robustness under simultaneous perturbations of all the variables together.

Analogous to that, we have the notion of *separate differentiation* and *joint differentiation*. Partial derivatives corresponds to the notion of *separate differentiation*. The corresponding notion in joint differentiation is called the *total derivative*. The notion of total derivative, however, requires some knowledge/understanding of linear algebra, which we cannot currently assume. Hence, we restrict our attention/analysis to partial derivatives.

1.2. Definition of partial derivative.

For simplicity, we restrict formulations and notation to functions of two variables, where we denote the two input variables as x and y. The same ideas apply to functions of more variables.

For a function f with input variables (x, y), we define the *partial derivative* of f with respect to x at the point (a, b), denoted $f_x(a, b)$, as the number:

$$f_x(a, b) := \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

In other words, it is the derivative with respect to the *first coordinate* keeping the second coordinate value fixed at b. In other words, it is the value $g'(a)$ where $g(x) := f(x, b)$.

Simiarly, we define the partial derivative of f with respect to y at the point (a, b), denoted $f_y(a, b)$, as the number:

$$f_y(a,b) := \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}$$

In other words, it is the derivative with respect to the *second coordinate* keeping the first coordinate value fixed at a. In other words, it is the value $g'(b)$ where $g(y) := f(a,y)$.

Note that since the above are calcuating partial derivatives at a *fixed* point, they give actual numbers. We could, however, now make a and b variable, and relabel them x and y. In this case $f_x(x,y)$ and $f_y(x,y)$ are now *both* functions of *both* variables x and y. The formulas look as follows:

$$\begin{aligned} f_x(x,y) &:= \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h} \\ f_y(x,y) &:= \lim_{h \to 0} \frac{f(x,y+h) - f(x,y)}{h} \end{aligned}$$

1.3. **The expression and value of partial derivative depend on both variable values.** One possible misconception is that the partial derivative with respect to a particular variable depends only on that variable. This is *not true*. The expression for the partial derivative with respect to x potentially depends on both x and y. What this means is that the value of the partial derivative depends on the location of the point, even the *other* coordinate.

The exception is the case of *additively separable functions*. In other words, if we can write $F(x,y)$ as $f(x)+g(y)$ where f is a function of one variable and g is a function of one variable. Then, $F_x(x,y) = f'(x)$ and is independent of y and $F_y(x,y) = g'(y)$ and is independent of x.

1.4. **Numerical subscripts for partials.** For functions of two variables, we can use the letter x for the first input and the letter y for the second input. This does not naturally generalize to functions with more inputs. Hence, there is an alternate convention: we use the subscript i to denote partial derivative with respect to the i^{th} input coordinate. Thus, $f_1(x,y)$ stands for $f_x(x,y)$ and $f_2(x,y)$ stands for $f_y(x,y)$.

Note that subscripts are often used in other contexts, so just because you see a subscript being used, do not blindly assume that it refers to a partial derivative. Context is everything.

1.5. **Leibniz-like notation for partial derivatives.** Recall that the Leibniz notation for ordinary differentiation uses the d/dx operator. For partial differentiation, we replace the English letter d by a letter ∂, so $f_x(x,y)$ is also denoted as $\frac{\partial}{\partial x}(f(x,y))$ and $f_y(x,y)$ is also denoted as $\frac{\partial}{\partial y}(f(x,y))$. In particular, if $z = f(x,y)$, we can write these partial derivatives as $\partial z/\partial x$ and $\partial z/\partial y$ respectively.

1.6. **Rule for computing partial derivatives.** Partial derivatives are computed just like ordinary derivatives – we just treat all the other input variables as constant. So, for instance:

$$\frac{\partial}{\partial x}(x^2 + xy + y^2) = \frac{\partial}{\partial x}(x^2) + y\frac{\partial}{\partial x}(x) + 0 = 2x + y$$

1.7. **Partial versus implicit differentiation.** In single variable calculus, you came across a concept called *implicit differentiation*, for which we used the letter d. With implicit differentiation, we start with an expression that involves both x and y, and then differentiate with respect to x. *However, for implicit differentiation, we do not assume that y is a constant.* Rather, we assume that y is an unknown implicit function of x, so our final expression involves dy/dx, which we *do not convert to zero*.

Partial differentiation with respect to x is different in that it does not assume y to be dependent on x – rather it assumes y is a constant, and treats y as such. However, if we have already done a calculation of the implicit derivative of an expression $f(x,y)$ with respect to x, we can calculate the partial derivative by simply setting $dy/dx = 0$ wherever it appears in the expression for the implicit derivative.

For instance, under implicit differentiation:

$$\frac{d}{dx}\sin(x + y + y^2) = \left[1 + (1+2y)\frac{dy}{dx}\right]\cos(x + y + y^2)$$

To compute the partial derivative, we simply set $dy/dx = 0$ in the above, and get:

$$\frac{\partial}{\partial x}\sin(x+y+y^2) = \cos(x+y+y^2)$$

1.8. **Partial derivatives plus implicit differentiation.** In the previous subsection, we contrasted partial differentiation of $F(x,y)$ with respect to x and implicit differentiation. The partial derivative can be obtained from the implicit derivative by setting $dy/dx = 0$, i.e., assuming that the variables have no dependence on each other.

When we have more than two variables, however, we may combine the ideas of partial and implicit differentiation. For instance, we may have an expression $G(x,y,z) = 0$ to describe z as an *implicit* function of the variables x and y. We now want to determine the partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$. In order to do this, we start with:

$$G(x,y,z) = 0$$

To find $\partial z/\partial x$, we differentiate $G(x,y,z)$ partially with respect to x, assuming the following: (i) y is treated as a constant, (ii) z is treated as an implicit function of x and y, so its partial derivative with respect to x is denoted $\partial z/\partial x$.

For instance, consider z as an implicit function of x and y:

$$xy + \sin(xz) = \cos(y+z)$$

Doing implicit differentiation, we get:

$$y + \cos(xz)\left(x\frac{\partial z}{\partial x} + z\right) = -\sin(y+z)\frac{\partial z}{\partial x}$$

We can rearrange, collect terms for $\partial z/\partial x$, and get an expression for it in terms of the three variables x, y, and z.

2. Conceptual and geometric interpretations

2.1. **What's a partial derivative, conceptually?** A given output typically depends on multiple inputs. For instance, in classical microeconomic theory, the quantity of a commodity demanded by a household is considered a function of six types of variables: the unit price, the tastes and preferences of the household, the income/wealth of the household, the prices of substitute goods, the prices of complementary goods, and expectations regarding future prices. Each of these "types" of variables may itself comprise multiple variables (for instance, there may be many different complementary and substitute goods), so the actual quantity demanded may be modeled as a function of a much larger number of variables. We may wish to study the effect of *just one of these* variables on the quantity demanded, keeping the other variables fixed. This is known as *ceteris paribus* in economics. The partial derivative is a key tool in the study of this kind of relationship.

If you have studied classical microeconomics in a quantitative sense, you may have encountered the concept of *price elasticity*. The price elasticity of demand is sort of like the partial derivative of the quantity demanded with respect to the price. However, in order to achieve dimensionlessness, we do not simply take the partial derivative but instead divide the partial derivative by the quantity-price ratio, i.e., we take the partial derivative of the *logarithm* of the quantity demanded with respect to the *logarithm* of the unit price of the good, keeping all the other determinants of demand constant. In symbols:

$$\text{Price elasticity of demand} = \frac{\partial q/\partial p}{q/p} = \frac{\partial(\ln q)}{\partial(\ln p)}$$

The quantity as computed here turns out to be negative if the good satisfies the law of demand, and it is customary to take the absolute value when giving numerical values.

Similarly, we can define the *cross price elasticity* with respect to a complementary or substitute good as the derivative of the logarithm of quantity demanded for a particular good with respect to the unit price of a particular complementary or substitute good. There is also a related notion of *income elasticity of demand.*

Note that the price elasticity and cross price elasticity values depend not only on the price, but also on the values of other determinants of demand. In other words, if we want to know the value of the price elasticity of demand for bread at a particular price of bread, we *also* need to specify all the other values of determinants of demands and only then can we compute the price elasticity. This is the same observation we made earlier: *the value of the partial derivative with respect to one input depends on the values of all inputs at the point of evaluation.*

To take an example that makes this context clear: we know that the price elasticity of demand is fairly high at prices close to those of substitutes, because the substitution effect operates most strongly. At prices much higher than the price of the substitute, demand is unfiformly lower, and at prices much lower than the price of the substitute, demand is uniformly higher.

Thus, if we were to change the price of a substitute good, that would affect the price ranges for which price elasticity of demand is high.

The *extent* to which changes in the values of one input affect the partial derivative with respect to another variable can be captured using *second-order mixed partial derivatives*, something we will see in a little while.

Derivatives cut both ways! Many "laws of economics" such as the law of demand and law of supply basically make an assertion about the sign of a partial derivative. For instance, the *law of demand* states that, ceteris paribus, as the unit price falls, the quantity demanded rises (or at least stays the same). This can be reframed as saying that the partial derivative of quantity demanded with respect to price is negative (or non-positive).

Surprisingly, people who lack calculus skills often believe that derivatives cut one way. For instance, they may agree with the statement that "as the price rises, the quantity falls" but disagree with the statement "as the price falls, the quantity rises." For instance, people are more likely to agree with the statement "as the price of accidents rises, people will be more careful" rather than the equivalent statement "as the price of accidents falls, people will be less careful." Now, there are ways of finding elusive wisdom in such affronts to calculus (for instance, differing values of price elasticity across the demand curve, exceptional demand curves, or long-run technological progress/secular trends) but at face value, if you agree with one statement, you should agree with the other.

2.2. **What's a partial derivative, geometrically?** Consider a function $z = f(x, y)$ of two variables. Consider the graph of this function, which is the surface $z = f(x, y)$. At a point (a, b):

- The value $f_x(a, b)$ is the slope (z/x-type) of the tangent line to the curve we obtain by intersecting the surface with the plane $y = b$ (which is parallel to the xz-plane). This intersection is basically the graph of the function $x \mapsto f(x, b)$.
- The value $f_y(a, b)$ is the slope (z/y-type) of the tangent line to the curve we obtain by intersecting the surface with the plane $x = a$. This intersection is basically the graph of the function $y \mapsto f(a, y)$.

We will explore this geometric interpretation in the next lecture.

3. Higher partial derivatives

3.1. **Basic definitions.** Any of the partial derivatives of a function of n variables is itself a function of the same n variables, so it can be differentiated again. We use the following basic notation for second partial derivatives of $z = f(x, y)$:

$$
\begin{aligned}
(f_x)_x = f_{xx} = f_{11} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2} \\
(f_x)_y = f_{xy} = f_{12} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 z}{\partial y \partial x} \\
(f_y)_x = f_{yx} = f_{21} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 z}{\partial x \partial y} \\
(f_y)_y = f_{yy} = f_{22} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}
\end{aligned}
$$

These are all called *second order partial derivatives* or *second partial derivatives* or (in short) *second partials* because they involve two differentiations. The second partials f_{xy} and f_{yx} are termed *mixed partial derivatives* or *mixed partials.*

3.2. **Clairaut's theorem: equality of mixed partials.** Clairaut's theorem states that if f is defined around a point (a, b) (i.e., in an open disk containing the point (a, b)) and if the mixed partials f_{xy} and f_{yx} are both continuous around (a, b), then:

$$f_{xy} = f_{yx}$$

We will get a better conceptual feel of this in a short while.

3.3. **Additively and multiplicatively separable functions.** Suppose $F(x, y)$ can be written as a sum $f(x) + g(y)$, i.e., it is *additively separable* in terms of functions of x and y. Then, we have $F_x(x, y) = f'(x)$ and $F_y(x, y) = g'(y)$. Moreover:

- All pure higher order derivatives with respect to x are the corresponding ordinary derivatives of f.
- All pure higher order derivatives with respect to y are the corresponding ordinary derivatives of g.
- All mixed partial derivatives are zero.

Another case of interest is where $F(x, y)$ can be written as a product $f(x)g(y)$, i.e., it is *multiplicatively separable* in terms of functions of x and y. Then, we have $F_x(x, y) = f'(x)g(y)$, $F_y(x, y) = f(x)g'(y)$, and $F_{xy}(x, y) = f'(x)g'(y)$. More generally, any mixed partial in which there are a many x's and b many y's is $f^{(a)}(x)g^{(b)}(y)$.

3.4. **Higher partials.** We can take higher order partials for functions in two variables. For instance, the third order partials are f_{xxx}, f_{xxy}, f_{xyx}, f_{xyy}, f_{yxx}, f_{yxy}, f_{yyx}, and f_{yyy}. This is a total of 8 possibilities. Clairaut's theorem, however, can be used to show that if the partials are all continuous, then what matters is only the number of x's and the number of y's – the sequence of differentiation does not matter. After this identification, there are four partials: f_{xxx}, f_{xxy}, f_{xyy}, and f_{yyy}.

3.5. **Functions of more than two variables.** The same notation and ideas apply. In particular, Clairaut's theorem reveals that, if the mixed partials are continuous around a point, then mixed partials that involve differentiation in each variable the same number of times must be equal. Thus, for instance, if $w = f(x, y, z)$ has all sixth order mixed partials continuous, then $f_{xxxyyz} = f_{yxyxzx}$.

3.6. **First versus higher partials: invariance under monotone transformations.** The first partial derivatives satisfy an important property: the sign of the first partial derivative is invariant under monotone transformations of the variables, such as replacing a variable by its logarithm, or its square (for positive variables) or its cube, etc. This means that when we say that A is increasing relative to B, the statement would still be true if we replaced A by $\ln A$ and B by B^3. In other words, the signs of the first partials depend only on the *ordinal scale* and not on the actual distance ratios involved.

On the other hand, the sign of the second and higher partials is not invariant under monotone transformations. If we replace the output function by the logarithm of the output function, the sign of the mixed partial may go from negative to positive or to zero. *While first partials do not depend on the choice of a correct measurement scale, higher partials are highly sensitive to the choice of measurement scale.*

4. Real world applications

4.1. **Higher partials, conceptually.** The mixed partial derivative of a function is an extremely important idea. It measures the sensitivity of the nature of how the function changes with respect to one variable as we change another variable. Let's consider some examples.

Suppose that, in order to achieve some output, you need two types of inputs: labor L and capital K. Your output is given by a function $f(L, K)$. The *marginal productivity* of labor is defined (roughly) as the partial derivative of output with respect to labor, i.e., the partial derivative $f_L(L, K)$ (again, for some definitions, we may choose to take logarithms before taking partial derivatives in order to obtain a dimensionless quantity, however, here for convenience we do not use logarithms). The marginal productivity of capital is defined as the partial derivative $f_K(L, K)$. Marginal productivity basically answers the question: if I increase the given factor of production ever so slightly, then to what extent is output affected? Or equivalently, if I decrease the given factor of production ever so slightly, then to what extent is output affected?

We now note the significance of the three second partials:

- The second partial f_{LL} measures whether the marginal product of labor is increasing or decreasing, and by how much. If $f_{LL} > 0$, that implies that labor is subject to *increasing returns*. This means that the more labor you put in, the more attractive it is to put in each additional unit of labor. On the other hand, if $f_{LL} < 0$, that implies that labor is subject to *diminishing returns*, i.e., the gains from adding additional units of labor becomes less (though they may still be positive) as we add more labor. Generally, we see that for a fixed value of K, labor is eventually subjected to diminishing returns, and possibly eventually even negative returns.

  For instance, if each worker sitting on a machine can produce 2 units (but workers take coffee breaks, so machines can be used to a slight extent by more than one worker), and if there are 90 computers, increasing the number of workers from 80 to 90 might increase output from 160 units to 180 units, but increasing the number of workers from 90 to 100 may produce very little increase, because the only way to squeeze more output is to have some workers use the machines while the others are on coffee break. This may, for instance, boost output only to 185 units. In this case, we see diminishing returns once the capital utilization starts getting complete.

  If $f_{LL} = 0$, we talk of constant returns to labor.
- We can similarly study f_{KK}, and talk of increasing returns to capital, decreasing returns to capital, and constant returns to capital.
- Finally, consider the mixed partial derivative $f_{LK} = f_{KL}$ (under the assumptions of Clairaut's theorem). This is basically addressing the question: how does the marginal productivity of labor get affected if we add more capital? Or put another way, how does the marginal productivity of capital get affected if we add more labor?

  If $f_{LK} > 0$, we say that *labor and capital are complementary inputs*. This is the typical situation, and it is not hard to think of examples. For instance, adding more machines increases the number of people who can be productively employed to operate the machines. This is particularly true if existing capital is close to being fully exploited. Put another way, adding more people makes it more worthwhile to obtain new machines for these people to operate.

  If $f_{LK} < 0$, we say that *labor and capital substitute for each other*. This means that adding more workers *reduces* the marginal product of capital. To think of this kind of situation, imagine that some of the tasks can be performed *either* by people or by machines. For instance, buying a machine that can generate calculus lectures reduces the need for, and the marginal product of, hiring a calculus lecturer. Or, buying a machine with payroll software reduces the amount of labor that the payroll manager needs to put in, and perhaps renders his job redundant.

Recall from what we learned of functions of one variable that increasing returns means the graph studying dependency only on that variable is concave up, constant returns means the graph studying dependency only on that variable is linear, and decreasing returns means the graph studying dependency only on that variable is concave down.

Note that the expressions f_{LL}, f_{LK}, and f_{KK} are not fixed numbers – they are themselves functions of L and K. In particular, this means that they need not have constant signs – the sign of f_{LL} may be positive for some values of (L, K) but negative for others. Similarly, the value of f_{LK} may be positive for some values of (L, K) and negative for others.

Interestingly, in the short run, the story is largely one of diminishing returns and a mix of complementary and substitution effects. In the somewhat longer run, on the other hand, the story is one of increasing returns and complementary effects. The chief reason is that over the longer run, it is possible to reconfigure the modes of production (through technological innovation, both low-tech and hi-tech) in order to better exploit the synergies between different resources. In the short run, for instance, machines may put people out of work by substituting for them, but in the longer run, people acquire new skills that complement those of the machines.[1]

Clairaut's theorem and intuition. Clairaut's theorem, although not counter-intuitive, is not entirely intuitive either. However, to get a good understanding of the interaction of multiple variables, you should try to make this a part of your intution.

[1]For instance, farmers put out of work due to greater mechanization of agriculture may end up having to settle for being computer engineers.

In our context, for instance, it says that the effect on the marginal product of labor of adding one unit of capital is the same as the effect on the marginal product of capital of adding one unit of labor. In other words, the extent to which labor makes capital more (or less) valuable is the same as the extent to which capital makes labor more (or less) valuable.

4.2. **A marriage of derivatives: like begets like, or opposites attract?** Turning from production to personal life, let's consider the question of marriage. Simplifying from the realities of the messy world, assume that there are two sexes (male and female), each person from one sex wants to marry a person from the other sex, and there is an equal number of people of each sex. Each person has a "quality score" and all males can be ranked by quality score, while all females can be ranked by quality score. A marriage creates a "household" whose goal is to maximize some kind of domestic production, which is a function of the quality scores of the two partners being married. If (following chromosome conventions) we denote the female's quality score by x and the male's quality score by y, then this is a function $f(x, y)$, and is increasing in both x and y.

Question: What way of matching males and females maximizes production? There are two extreme possibilities. The first is *assortative mating*, where the highest quality males join hands with the highest quality females, and the lowest quality males walk the aisle with the lowest quality females. We might call this *like begets like theory of mating*. The other is the reverse, where the highest quality males marry the lowest quality females, and the lowest quality males marry the highest quality females. We might call this the *opposites attract theory of mating*.

Which story maximizes domestic production depends on *mixed partials*. Specifically, if the domestic production function has a positive mixed partial with respect to male and female quality, i.e., if $f_{xy} > 0$, that means that the higher the male quality, the *more* beneficial it is to marry a female of higher quality. Similarly, the higher the female quality, the more beneficial it is to marry a male of higher quality. In other words, *a positive mixed partial derivative bodes well for assortative mating*.

On the other hand, if $f_{xy} < 0$, then the higher the male quality, the less beneficial it is to marry a female of higher quality. This is a subtle point so it's worth pondering a bit. For any given male, if he has the choice, it always makes sense to marry the highest quality female he can get. But with a negative mixed derivative, the *margin of difference between high quality females and low quality females* is *lower for high quality males*. On the other hand, low quality males see a significant boost from attracting high quality females. Thus, the utility-maximizing arrangement would be one where the lowest quality males pair up with the highest quality females, and the highest quality males pair up with the lowest quality females.

One could also set up a bidding/auction scenario to see how, under an open bidding process for mates, this utility-maximizing outcome can be achieved. Suppose females are bidding for males. Basically, in the positive mixed partial derivative case, the high quality females are more desperate for the high quality males than the low quality females are, so they outbid the low quality females, leaving the low quality females to make do with the low quality males. On the other hand, with negative mixed partials, the low quality females, despite being less attractive than high quality females, are still able to get the high quality males because they are more desperate to win over the males and are willing to offer better terms. The disgruntled high quality females settle for the low quality males left in the pool, who are glad at their catch.

For what it's worth, most of the evidence suggests that mating is highly assortative. This doesn't quite prove that domestic production enjoys positive mixed partials, but it suggests that if we are trying to fit the real world into our highly restrictive model, then positive mixed partials would generate more realistic predictions.

4.3. **Cobb-Douglas production function.** The Cobb-Douglas production function is a particular form of a production function describing output in terms of two or more inputs. We consider the case of two inputs, labor L and capital K. Suppose the output is given by a function:

$$f(L, K) := CL^a K^b$$

Note that the numbers L and K denote the financial expenditures on labor and capital.

Here, C, a, and b are all positive and L and K are restricted to positive inputs. Logarithmically, we get:

$$\ln(f(L, K)) = \ln C + a \ln L + b \ln K$$

The partial derivatives are as follows:

$$\begin{aligned}
\frac{\partial}{\partial L} f(L,K) &= CaL^{a-1}K^b \\
\frac{\partial}{\partial K} f(L,K) &= CbL^aK^{b-1} \\
\frac{\partial^2}{\partial L^2} f(L,K) &= Ca(a-1)L^{a-2}K^b \\
\frac{\partial^2}{\partial K^2} f(L,K) &= Cb(b-1)L^aK^{b-2} \\
\frac{\partial^2}{\partial L \partial K}(f(L,K)) &= CabL^{a-1}K^{b-1} \\
\frac{\partial}{\partial(\ln L)}(\ln(f(L,K))) &= a \\
\frac{\partial}{\partial(\ln K)}(\ln(f(L,K))) &= b \\
\frac{\partial^2}{\partial(\ln L)\partial(\ln K)}(\ln(f(L,K))) &= 0
\end{aligned}$$

The partial derivatives involving logarithms are the dimensionless versions of the parital derivatives, in the same sense that price elasticity is a dimensionless version of the partial derivative of quantity demanded with respect to price.

4.4. **Returns to scale.** We note that:

- If labor alone is multiplied by a factor of λ, then output gets multiplied by a factor of λ^a. From this, or by looking at the second derivative, we see that if $0 < a < 1$ we have positive but decreasing returns on labor holding capital fixed. If $a = 1$, we have constant returns on labor holding capital fixed. If $a > 1$, we have positive returns on labor holding capital fixed.
- If capital alone is multiplied by a factor of λ, then output gets multiplied by a factor of λ^b. From this, or by looking at the second derivative, we see that if $0 < b < 1$, then there are positive but decreasing returns on capital holding labor fixed. For $b = 1$, constant returns on capital holding labor fixed, and for $b > 1$, increasing returns on capital holding labor fixed.
- If we multiply labor and capital both simultaneously by the same factor of λ, output is multiplied by a factor of λ^{a+b}. Thus, we get decreasing, constant, or increasing returns to scale depending on whether $a + b < 1$, $a + b = 1$, or $a + b > 1$.
- Note that in any Cobb-Douglas production model, labor and capital always play complementary roles, because the mixed partial derivative is positive.

4.5. **Maximizing production: fixed total of labor and capital.** If the total investment in labor and capital is fixed, i.e., $L + K$ is fixed, then, in order to maximize the production, we must allocate resources of L and K in the proportion $a : b$. In other words, if $L + K = M$ for some constant M, then the maximum production occurs when $L = Ma/(a + b)$ and $K = Mb/(a + b)$.

To see this, note that if x is the fraction allocated to labor, and $1 - x$ to capital, then the output is $CM^{a+b}x^a(1 - x)^b$. Differentiating with respect to x and then computing the local maximum by setting the derivative to zero gives the answer.

Roughly speaking, this means that the proportion of investment in labor should be based on the exponents on labor and capital. To determine whether investment is being done the way it "should" we need to find out:

- The exponents a and b in the Cobb-Douglas production function.
- The proportion of investment used for labor.

Historically, this was the motivation for the creation of the Cobb-Douglas production function. Douglas assembled historical data and estimated that production could be modeled roughly by a Cobb-Douglas production function with the exponents a and b in the ratio 3 : 1. He also found that the expenditures on labor and capital were in the ratio 3 : 1. Thus, he concluded that things were being done the way they "should" be.

Although a production function need not be a Cobb-Douglas production function, trying to use a Cobb-Douglas model is a good first pass.

5. Partial differential equations

This is a fairly tricky topic and we consider it in a very superficial way.

5.1. **Ordinary differential equations: reminder.** Recall the notion of (*ordinary*) differential equation. Here, there is one dependent variable (typically denoted y) and one independent variable (typically denoted x). The ordinary differential equation is of the form:

$$F(x, y, y', y'', \dots, y^{(k)}) = 0$$

In other words, some expression involving x, y, and the derivatives (first and higher) of y with respect to x is zero. A "solution" to this equation is a function (given either explicitly or via an implicit/relational description) $y = f(x)$ such that the above equation holds *identically* for all x in the domain.

Note that not every equation involving derivatives of functions of one variable is a differential equation. For instance, the equation $f'(x) = f(1-x)f(2-x)$ is a functional equation involving derivatives, but it is *not* a differential equation because it involves evaluating the function f at multiple points. A differential equation, as we use the term, refers to a situation that is restricted to the local behavior around a *single* point.

5.2. **Partial differential equations.** Partial differential equations are to functions of many variables what ordinary differential equations are to functions of one variable. Specifically, a partial differential equation involves some expression in terms of a bunch of variables, a function of those variables, and partial derivatives of that function.

Just like the case of ordinary differential equations, it remains true that a partial differential equation must involve *only* the behavior at and around a particular point in the domain. For instance, if we have a function $u = f(x, y)$ then the equation $f(x, y) = f_y(x^2, x + y) - f_{xx}(x^2, y^2)$ is a functional equation involving derivatives but it is *not* a partial differential equation.

The *order* of a partial differential equation is defined as the highest of the orders of the partial derivatives that appear in the partial differential equation.

5.3. **Key examples of partial differential equations.** The *Laplace equation* is an example of a partial differential equation. For a function u of variables x and y, the equation is:

$$u_{xx} + u_{yy} = 0$$

In Leibniz notation, this becomes:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

Another example is the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$$

5.4. **How big is the space of solutions functions?** Recall that for an ordinary differential equation, the solution function need not be unique, but the space of solution functions can typically be described as a r-dimensional space (in the sense of there being r free parameters) where r is the order of the differential equation.

For a partial differential equation, the solution space is *much much larger*. Usually it is infinite-dimensional if we are dealing with functions of more than one variable. It is usually impossible to give a description of the general solution.