

Linear and quadratic Taylor polynomials for functions of several variables.

© 2018, Yonatan Katznelson

Finding the extreme (minimum or maximum) values of a function, is one of the most important applications of differential calculus to economics. In general, there are two steps to this process: (i) finding the points where extreme values may occur, and (ii) analyzing the behavior of the function near these points, to determine whether or not extreme values actually occur there. For a function of one variable, we saw that step (i) consisted of finding the critical points of the function, and step (ii) consisted of using the first or second derivative test to analyze the behavior of the function near the point.

We use the same two steps to find the critical points and classify the critical values of functions of several variables, and to understand how the procedure generalizes, we need to first understand the linear and quadratic Taylor's polynomials for functions of several variables, and the approximations that they provide.

Comment: I will assume throughout this note that all the functions being discussed have continuous first, second and third order derivatives (or partial derivatives). This assumption justifies the claims made below about the accuracy of the approximations.

1. The one-variable case.

To make sense of Taylor polynomials in several variables, we first recall what they look like for functions of one variable.[†]

1.1 The first order (linear) approximation

The first order Taylor polynomial for $y = f(t)$, centered at t_0 is the *linear* function

$$T_1(t) = f(t_0) + f'(t_0) \cdot (t - t_0). \quad (1)$$

With this definition of T_1 , the approximation

$$f(t) \approx T_1(t), \quad (2)$$

is very good if t is sufficiently close to t_0 .[‡]

This approximation is also called the *tangent line approximation* because the graph $y = T_1(t)$ is the tangent line to the graph $y = f(t)$ at the point $(t_0, f(t_0))$.

1.2 The second order (quadratic) approximation

The second order Taylor polynomial for the function $y = f(t)$, centered at t_0 is the *quadratic* function

$$T_2(t) = f(t_0) + f'(t_0)(t - t_0) + \frac{f''(t_0)}{2}(t - t_0)^2. \quad (3)$$

[†]For a more thorough discussion of Taylor polynomials for functions of one variable, please see **SN 7** on the review page of the 11A website: <http://people.ucsc.edu/~yorik/11A/review.htm>.

[‡]The error of approximation $|f(t) - T_1(t)|$ is less than a multiple of $|t - t_0|^2$. As t approaches t_0 , the squared difference $|t - t_0|^2$ goes to 0 very rapidly.

You should note that $T_2(t) = T_1(t) + \frac{f''(t_0)}{2}(t-t_0)^2$. As for the linear approximation, the quadratic approximation

$$f(t) \approx T_2(t). \quad (4)$$

is very good if t is sufficiently close to t_0 . In fact, the quadratic approximation is typically much better than the linear approximation once t is very close to t_0 .[§]

At this point the key question is:

Why do the Taylor polynomials $T_1(t)$ and $T_2(t)$ do such a good job of approximating the original function $f(t)$ in the neighborhood of t_0 ?

The answer is provided by looking at what happens at the point t_0 itself. The function $T_1(t)$ is the unique linear function that satisfies the conditions

$$T_1(t_0) = f(t_0) \quad \text{and} \quad T'_1(t_0) = f'(t_0). \quad (5)$$

That is, $T_1(t)$ has the same value as $f(t)$ at t_0 and $T_1(t)$ is changing at the same rate as $f(t)$ at t_0 .

Likewise, the function $T_2(t)$ is the unique quadratic function that has the same value at t_0 as $f(t)$, has the same rate of change at t_0 as $f(t)$ and has the same *concavity* at t_0 as $f(t)$ ($T'_2(t)$ is changing at the same rate as $f'(t)$ at the point t_0). I.e.,

$$T_2(t_0) = f(t_0), \quad T'_2(t_0) = f'(t_0) \quad \text{and} \quad T''_2(t_0) = f''(t_0). \quad (6)$$

In other words, the functions $T_1(t)$ and $T_2(t)$ behave very much like $f(t)$ at the point t_0 , and this similar behavior *at* t_0 carries over to points that are *close* to t_0 .

To find analogous polynomial approximations for a function of several variables, we will impose the same sort of conditions on the values of the polynomial and on the values of the *partial derivatives* of the polynomial at a specified point. This will give simple equations for the coefficients of the polynomial, as I will describe in the sections that follow.

2. The linear Taylor polynomial in two variables.

Suppose that we want to find a linear polynomial in two variables that approximates the (differentiable) function $F(x, y)$ in the neighborhood of the point (x_0, y_0) . That is, we want to find a function

$$T_1(x, y) = A + B(x - x_0) + C(y - y_0)$$

such that

$$T_1(x, y) \approx F(x, y)$$

if (x, y) is close to (x_0, y_0) . We choose to write $T_1(x, y)$ as a function of $(x - x_0)$ and $(y - y_0)$ (instead of the more traditional $T_1(x, y) = a + bx + cy$) because it emphasizes the role of the point (x_0, y_0) in the approximation and it makes the equations we will solve *much* simpler.

The comments at the end of the previous section lead to the idea of imposing conditions on the values of $T_1(x, y)$, $T_{1x}(x, y)$ and $T_{1y}(x, y)$ at the point (x_0, y_0) . In analogy with the properties of $T_1(t)$ described in (5), we require that

$$T_1(x_0, y_0) = F(x_0, y_0), \quad \frac{\partial T_1}{\partial x}(x_0, y_0) = \frac{\partial F}{\partial x}(x_0, y_0) \quad \text{and} \quad \frac{\partial T_1}{\partial y}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0).$$

[§]If $|t - t_0|$ is small, then $|f(t) - T_2(t)|$ is less than a multiple of $|t - t_0|^3$. Thus, if $|t - t_0| < 1$, then both $|t - t_0|^2$ and $|t - t_0|^3$ are *very* small and $|t - t_0|^3 < |t - t_0^2|$. This implies (with more work) that once $|t - t_0|$ is small enough, $|f(t) - T_2(t)| < |f(t) - T_1(t)|$.

These conditions give *very* simple equations for the unknown coefficients A , B and C .

Specifically, since $T_1(x_0, y_0) = A$, it follows that $A = F(x_0, y_0)$; since $T_{1x} = B$, it follows that $B = F_x(x_0, y_0)$; and since $T_{1y} = C$, it follows that $C = F_y(x_0, y_0)$. In other words, the first order Taylor polynomial for $F(x, y)$, centered at (x_0, y_0) is given by

$$T_1(x, y) = F(x_0, y_0) + F_x(x_0, y_0) \cdot (x - x_0) + F_y(x_0, y_0) \cdot (y - y_0). \quad (7)$$

Using more advanced techniques,[¶] it possible to show that the approximation

$$F(x, y) \approx T_1(x, y)$$

is very good if (x, y) is close to (x_0, y_0) .

Example 1. If $F(x, y) = \sqrt{5x + 2y}$ and $(x_0, y_0) = (1, 10)$, then

$$F_x = \frac{5}{2}(5x + 2y)^{-1/2} \quad \text{and} \quad F_y = (5x + 2y)^{-1/2},$$

so

$$F(1, 10) = \sqrt{25} = 5, \quad F_x(1, 10) = \frac{5}{10} = \frac{1}{2} \quad \text{and} \quad F_y(1, 10) = \frac{1}{5}.$$

It follows that the linear Taylor approximation to $\sqrt{5x + 2y}$, centered at $(1, 10)$, is given by

$$T_1(x, y) = 5 + \frac{1}{2}(x - 1) + \frac{1}{5}(y - 10).$$

To ‘test’ the accuracy of the approximation that this provides, I’ll use it to estimate

$$\sqrt{27.04} = \sqrt{5 \cdot (1.2) + 2 \cdot (10.52)} = F(1.2, 10.52).$$

Plugging $(x, y) = (1.2, 10.52)$ into $T_1(x, y)$, we have

$$T_1(1.2, 10.52) = 5 + \frac{0.2}{2} + \frac{0.52}{5} = 5.204,$$

giving the estimate

$$\sqrt{27.04} = F(1.2, 10.52) \approx T_1(1.2, 10.52) = 5.204.$$

This estimate is exactly $1/250$ away from the truth, since $\sqrt{27.04} = 5.2$ and $0.004 = 1/250$.

3. The quadratic Taylor polynomial in two variables.

To find the formula of the quadratic Taylor approximation for the function $F(x, y)$, centered at the point (x_0, y_0) , we repeat the procedure we followed above for the linear polynomial, but we take it one step further.

In analogy with the conditions satisfied by $T_2(t)$ in the one-variable setting (shown in (6)), we want to find the coefficients A, B, C, D, E and G of the quadratic function

$$T_2(x, y) = A + B(x - x_0) + C(y - y_0) + D(x - x_0)^2 + E(y - y_0)^2 + G(x - x_0)(y - y_0)$$

[¶]These techniques are outside the scope of this course, but well within the realm of an upper division math class.

that satisfies the conditions

$$\begin{aligned}
T_2(x_0, y_0) &= F(x_0, y_0), \\
\frac{\partial T_2}{\partial x}(x_0, y_0) &= \frac{\partial F}{\partial x}(x_0, y_0), \\
\frac{\partial T_2}{\partial y}(x_0, y_0) &= \frac{\partial F}{\partial y}(x_0, y_0), \\
\frac{\partial^2 T_2}{\partial x^2}(x_0, y_0) &= \frac{\partial^2 F}{\partial x^2}(x_0, y_0), \\
\frac{\partial^2 T_2}{\partial y^2}(x_0, y_0) &= \frac{\partial^2 F}{\partial y^2}(x_0, y_0), \\
\frac{\partial^2 T_2}{\partial x \partial y}(x_0, y_0) &= \frac{\partial^2 F}{\partial x \partial y}(x_0, y_0).
\end{aligned}$$

In other words, we want $T_2(x, y)$, its first order partial derivatives and its second order partial derivatives to *all* have the same values as the corresponding derivatives of $F(x, y)$ at the point (x_0, y_0) .

The values of the partial derivatives of $T_2(x, y)$ at the point (x_0, y_0) are all very simple expressions in the coefficients of $T_2(x, y)$. Indeed, we have

$$\begin{aligned}
T_2(x_0, y_0) &= A, \\
\frac{\partial T_2}{\partial x}(x_0, y_0) &= B, \\
\frac{\partial T_2}{\partial y}(x_0, y_0) &= C, \\
\frac{\partial^2 T_2}{\partial x^2}(x_0, y_0) &= 2D, \\
\frac{\partial^2 T_2}{\partial y^2}(x_0, y_0) &= 2E, \\
\frac{\partial^2 T_2}{\partial x \partial y}(x_0, y_0) &= G,
\end{aligned}$$

as you should verify by doing the calculations yourself.

Comparing the two lists above, we find that the quadratic Taylor polynomial for $F(x, y)$, centered at (x_0, y_0) , is given by

$$\begin{aligned}
T_2(x, y) &= F(x_0, y_0) + F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0) \\
&\quad + \frac{F_{xx}(x_0, y_0)}{2}(x - x_0)^2 + \frac{F_{yy}(x_0, y_0)}{2}(y - y_0)^2 + F_{xy}(x_0, y_0)(x - x_0)(y - y_0).
\end{aligned} \tag{8}$$

As in the one-variable case, the linear and constant coefficients of $T_2(x, y)$ are the same as those of $T_1(x, y)$. In other words, we have

$$T_2(x, y) = T_1(x, y) + \frac{F_{xx}(x_0, y_0)}{2}(x - x_0)^2 + \frac{F_{yy}(x_0, y_0)}{2}(y - y_0)^2 + F_{xy}(x_0, y_0)(x - x_0)(y - y_0).$$

Also as in the one-variable case, the quadratic terms in $T_2(x, y)$ tend to make the approximation $F(x, y) \approx T_2(x, y)$ more accurate than the approximation $F(x, y) \approx T_1(x, y)$, when (x, y) is very close to (x_0, y_0) .

Example 2. Let's see if the quadratic Taylor polynomial gives a more accurate approximation than the linear one for the function and points in Example 1. We already have the constant and linear coefficients of T_2 (why?), so it remains to find the quadratic coefficients.

First we compute the second order partial derivatives of $F(x, y) = \sqrt{5x + 2y}$:

$$F_{xx} = -\frac{25}{4}(5x + 2y)^{-3/2}, \quad F_{yy} = -(5x + 2y)^{-3/2} \quad \text{and} \quad F_{xy} = -\frac{5}{2}(5x + 2y)^{-3/2}.$$

Next, we evaluate these at the point $(x_0, y_0) = (1, 10)$, remembering that $25^{-3/2} = \frac{1}{125}$:

$$F_{xx}(1, 10) = -\frac{1}{20}, \quad F_{yy}(1, 10) = -\frac{1}{125} \quad \text{and} \quad F_{xy}(1, 10) = -\frac{1}{50}.$$

Using these numbers and the ones we found in Example 1 in Equation (8), we find that

$$T_2(x, y) = 5 + \frac{1}{2}(x - 1) + \frac{1}{5}(y - 10) - \frac{1}{40}(x - 1)^2 - \frac{1}{250}(y - 10)^2 - \frac{1}{50}(x - 1)(y - 10),$$

and

$$T_2(1.2, 10.52) = 5 + \frac{0.2}{2} + \frac{0.52}{5} - \frac{(0.2)^2}{40} - \frac{(0.52)^2}{250} - \frac{(0.2)(0.52)}{50} = 5.1998384\dots$$

This give the estimate

$$\sqrt{27.04} = F(1.2, 10.52) \approx T_2(1.2, 10.52) = 5.1998384\dots,$$

which is closer to the correct value ($\sqrt{27.04} = 5.2$) than the previous estimate, since

$$|T_2(1.2, 10.52) - F(1.2, 10.52)| = 0.000161616\dots \left(= \frac{101}{625000} \right).$$

This is almost 25 times smaller than the error of approximation we had when we used the linear approximation.

4. Linear and quadratic Taylor polynomials for functions of three variables.

As the number of variables grows, the number of terms in both the linear and quadratic Taylor polynomials grows as well, but the forms of the polynomials follow the same patterns that we saw in (7) and (8) for the two-variable case. The formulas for a function of three variables appear below.

4.1 The linear Taylor polynomial for a function of three variables.

The linear Taylor polynomial for the function $F(x, y, z)$, centered at the point (x_0, y_0, z_0) , has the form

$$T_1(x, y, z) = F(x_0, y_0, z_0) + F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0)$$

(9)

As in the one- and two-variable cases, $T_1(x_1, \dots, x_n)$ satisfies the conditions

$$\begin{aligned} T_1(x_0, y_0, z_0) &= F(x_0, y_0, z_0), \\ T_{1x}(x_0, y_0, z_0) &= F_x(x_0, y_0, z_0), \\ T_{1y}(x_0, y_0, z_0) &= F_y(x_0, y_0, z_0), \\ T_{1z}(x_0, y_0, z_0) &= F_z(x_0, y_0, z_0). \end{aligned}$$

Indeed, as before, it is by requiring that these conditions be satisfied that the coefficients were found. Also as before, the approximation

$$F(x, y, z) \approx T_1(x, y, z)$$

is very accurate when (x, y, z) is sufficiently close to (x_0, y_0, z_0) .

4.2 The quadratic Taylor polynomial for a function of three variables.

The linear Taylor polynomial for the function $F(x, y, z)$, centered at the point (x_0, y_0, z_0) , has the form

$$\begin{aligned} T_2(x, y, z) = & F(x_0, y_0, z_0) + F_x(x_0, y_0, z_0)(x - x_0) \\ & + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) \\ & + \frac{F_{xx}(x_0, y_0, z_0)}{2}(x - x_0)^2 + \frac{F_{yy}(x_0, y_0, z_0)}{2}(y - y_0)^2 \\ & + \frac{F_{zz}(x_0, y_0, z_0)}{2}(z - z_0)^2 + F_{xy}(x_0, y_0, z_0)(x - x_0)(y - y_0) \\ & + F_{xz}(x_0, y_0, z_0)(x - x_0)(z - z_0) + F_{yz}(x_0, y_0, z_0)(y - y_0)(z - z_0) \end{aligned} \quad (10)$$

The coefficients of $T_2(x, y, z)$ are chosen so that the value of T_2 and those of all of its first and second order partial derivatives are equal to the values of the corresponding partial derivatives of $F(x, y, z)$. These conditions ensure, as before, that the approximation

$$F(x, y, z) \approx T(x, y, z)$$

is very accurate if (x, y, z) is sufficiently close to (x_0, y_0, z_0) .

5. Another example.

The example in this section is meant to illustrate how the approximation a function of two variables by its quadratic Taylor polynomial is used in the classification of local extreme values. In particular you will see that we don't use the approximation to estimate specific values, but rather we use the Taylor polynomial to approximate the overall behavior of the function in the neighborhood of a critical point. This idea eventually leads to the generalization of the second derivative test to the case of functions of two or more variables.

Consider the function

$$f(x, y) = \left(1 - \frac{x^3}{2} - \frac{y^3}{3}\right) e^{0.75x^2+y^2},$$

whose graph in the neighborhood of $(0, 0)$ appears in Figure 1, below. The point $(0, 0, 1)$ on the graph is the point from which the vertical arrow is extending, and it appears to be the lowest point on the graph in its own immediate vicinity. In other words, it appears that $f(0, 0) = 1$ is a *local minimum* value of the function, meaning that

$$f(0, 0) < f(x, y)$$

for all points (x, y) that are sufficiently close to $(0, 0)$. However pictures can be misleading, so it is important to have an analytic argument that *proves* that $f(0, 0)$ is indeed a local minimum value.

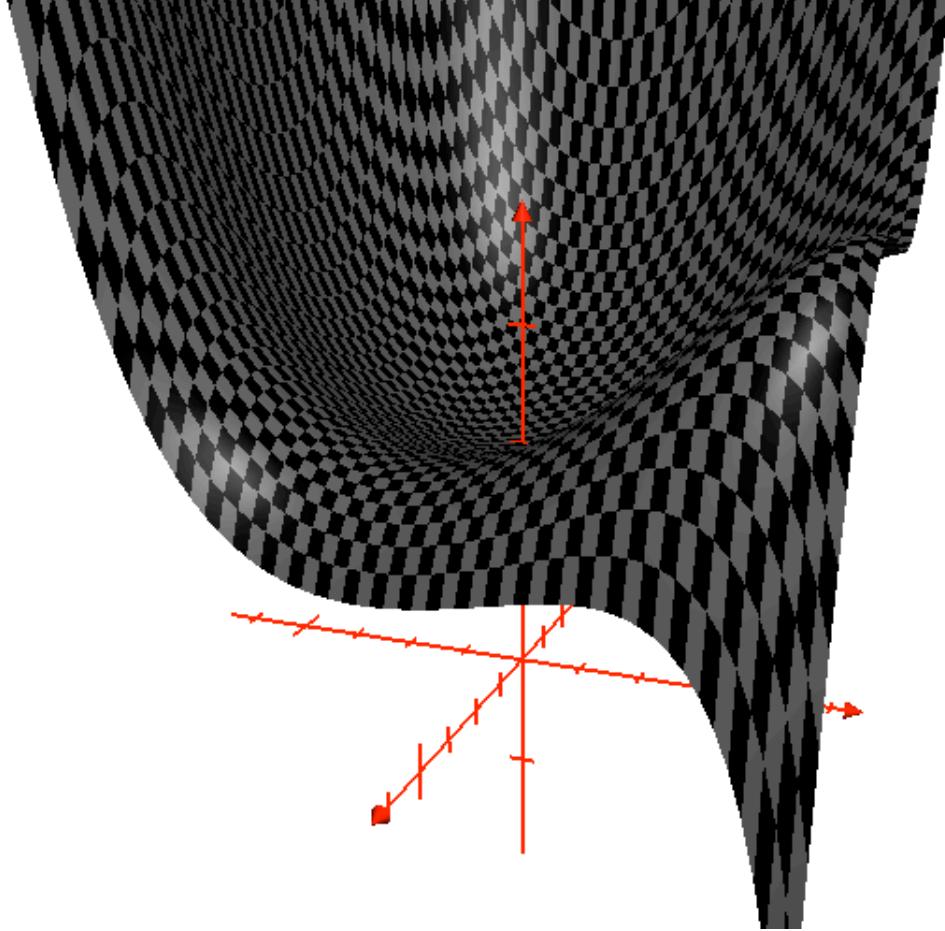


Figure 1: The graph of $f(x, y) = \left(1 - \frac{x^3}{2} - \frac{y^3}{3}\right) e^{0.75x^2+y^2}$.

Analyzing the function $f(x, y)$ as it is defined may be a little complicated, and this is where the Taylor approximation of $f(x, y)$, centered at $(0, 0)$, becomes very useful. Applying the Taylor approximation, $f(x, y) \approx T_2(x, y)$, in this case, yields

$$\begin{aligned}
f(x, y) &\approx 1 + f_x(0, 0)(x - 0) + f_y(0, 0)(y - 0) + \frac{f_{xx}(0, 0)}{2}(x - 0)^2 \\
&\quad + \frac{f_{yy}(0, 0)}{2}(y - 0)^2 + f_{xy}(0, 0)(x - 0)(y - 0) \\
&= 1 + f_x(0, 0) \cdot x + f_y(0, 0) \cdot y + \frac{f_{xx}(0, 0)}{2} \cdot x^2 + \frac{f_{yy}(0, 0)}{2} \cdot y^2 + f_{xy}(0, 0) \cdot xy,
\end{aligned} \tag{11}$$

(which is accurate if (x, y) is sufficiently close to $(0, 0)$). The expression in the third row is fairly simple already, and it will become simpler still once we evaluate all the partial derivatives of $f(x, y)$ at $(0, 0)$ and compute the (numerical) coefficients.

Differentiating once yields

$$f_x(x, y) = -\frac{3x^2}{2}e^{0.75x^2+y^2} + 1.5x \left(1 - \frac{x^3}{2} - \frac{y^3}{3}\right) e^{0.75x^2+y^2}$$

and

$$f_y(x, y) = -y^2e^{0.75x^2+y^2} + 2y \left(1 - \frac{x^3}{2} - \frac{y^3}{3}\right) e^{0.75x^2+y^2},$$

and evaluating these expressions at the point $(0, 0)$ gives

$$f_x(0, 0) = 0 \quad \text{and} \quad f_y(0, 0) = 0,$$

which means that $T_2(x, y)$ has *no linear terms* in this instance.^{||}

Differentiating the first derivatives (and collecting terms) yields the second order derivatives:

$$f_{xx}(x, y) = \left[- (3x + 4.5x^3) + (1.5 + 2.25x^2) \left(1 - \frac{x^3}{2} - \frac{y^3}{3} \right) \right] e^{0.75x^2+y^2},$$

$$f_{xy}(x, y) = \left[-3x^2y - 1.5xy^2 + 3xy \left(1 - \frac{x^3}{2} - \frac{y^3}{3} \right) \right] e^{0.75x^2+y^2}$$

and

$$f_{yy}(x, y) = \left[- (2y + 4y^3) + (2 + 4y^2) \left(1 - \frac{x^3}{2} - \frac{y^3}{3} \right) \right] e^{0.75x^2+y^2}.$$

As complicated as they appear,^{**} the second order partial derivatives are very easy to evaluate at the point $(0, 0)$. Indeed

$$f_{xx}(0, 0) = [-0 + (1.5 + 0)(1 - 0 - 0)] e^0 = 1.5 \cdot 1 \cdot 1 = 1.5,$$

$$f_{xy}(0, 0) = [0 - 0 + 0(1 - 0 - 0)] e^0 = 0$$

and

$$f_{yy}(0, 0) = [-0 + (2 + 0)(1 - 0 - 0)] e^0 = 2 \cdot 1 \cdot 1 = 2.$$

Returning to the Taylor approximation (11), and inserting the values of the coefficients, we see that if (x, y) is close to $(0, 0)$, then

$$f(x, y) \approx 1 + \frac{3x^2}{4} + y^2.$$

This means that if (x, y) is close (but not equal to) to $(0, 0)$, then

$$f(x, y) - f(0, 0) \approx \frac{3x^2}{4} + y^2 > 0.$$

In other words, if (x, y) is sufficiently close to $(0, 0)$, then $f(x, y) \geq f(0, 0)$, showing that $f(0, 0)$ is a local minimum value, as we suspected.

^{||}The conditions $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$ make (x_0, y_0) a *critical point* of the function $f(x, y)$, as you might expect. More on that in SN 3.

^{**}I recommend that you compute these partial derivatives on your own, for practice.

6. Exercises

1. Compute $T_2(x, y)$ for the function

$$f(x, y) = x^3y + 3x^2y^2 - 2xy^2 + 3x - 5y,$$

centered at the point $(x_0, y_0) = (1, 1)$.

2. Compute $T_2(K, L)$ for the function $Q(K, L) = 10K^{2/3}L^{1/3}$, centered at the point $(K_0, L_0) = (1, 8)$. Use the Taylor polynomial that you found to calculate the *approximate* value of $Q(1.5, 8.4)$.
3. Compute $T_2(u, v)$ for the function $g(u, v) = (2u + 3v)e^{u^2+v^2}$, centered at the point $(u_0, v_0) = (0, 0)$.
4. Compute $T_2(x, y)$ for the function $H(x, y) = \sqrt{2x + 5y}$, centered at the point $(x_0, y_0) = (3, 2)$. Use the Taylor polynomial that you found to calculate the approximate value of $H(3.25, 2.1) = \sqrt{17}$.
5. Compute $T_2(x, y, z)$ for the function

$$g(x, y, z) = 2 \ln x + 4 \ln y - z(5x + 8y - 60),$$

centered at the point $(x_0, y_0, z_0) = (4, 5, 0.1)$.