Part A

A.1

Cobb-Douglas: Round 2 (***#diffapplication***)

Task (a)

(i)
To find the plane tangent to $Y(L, K)$ at point (1000 hours, $1 million), we need to find the partial derivatives for variables L and K.

$$\frac{\partial Y(10^3, 10^6)}{\partial L} = 1.5\frac{K^{0.25}}{L^{0.25}} = 1.5 * 10^{0.75}$$
$$\frac{\partial Y(10^3, 10^6)}{\partial K} = 0.5\frac{L^{0.75}}{K^{0.75}} = \frac{0.5}{10^{2.25}}$$

Then, we plug these values into the equation of a tangent plane:

$$Y_{tan}(L, K) = Y(L_0, K_0) + \frac{\partial Y(10^3, 10^6)}{\partial L}(L - L_0) + \frac{\partial Y(10^3, 10^6)}{\partial K}(K - K_0)$$
$$Y_{tan}(L, K) = 10^{6.75} + 1.5 * 10^{0.75}(L - 10^3) + \frac{0.5}{10^{2.25}}(K - 10^6)$$

(ii)
To find the linear approximation at $Y(10^3 + 1, 10^6 + 1)$, we will plug these values in instead of L and K in the equation above:

$$Y_{tan}(10^3 + 1, 10^6 + 1) = 2 * 10^{3.75} + 1.5 * 10^{0.75} * 1 + \frac{0.5}{10^{2.25}} * 1 \approx 11255.2644$$

The change then would be:

$$Y_{tan}(10^3 + 1, 10^6 + 1) - Y(10^3, 10^6) \approx 8.4379$$

(iii)
The actual function value at point (10^3 + 1, 10^6 + 1) is:

$$Y(10^3 + 1, 10^6 + 1) = 2 * (10^3 + 1)^{0.75} * (10^6 + 1)^{0.25} \approx 11246.8265$$

Then, the actual change is:

$$Y(10^3 + 1, 10^6 + 1) - Y(10^3, 10^6) = 8.4368$$

Let's find the relative error:

$$\frac{|Y_{tan}(10^3 + 1, 10^6 + 1) - Y(10^3 + 1, 10^6 + 1)|}{Y(10^3 + 1, 10^6 + 1) - Y(10^3, 10^6)} \approx 0.0125$$

This relative error is very small, so we can count our approximation as very decent. It is quite expectable because the changes in 1 hour compared to 1000 hours and $1 compared to $1M are very tiny, so that new point is within the boundaries of (1000 h, $1M), where the linear approximation does a very good job.

Task (b)

Directional derivatives here show the ratio of change of real value of goods produced in a year (**Y**) when we vary (take into account different configurations of) the rates of change for person-hours worked in a year (**L**) and real capital change (**K**).

The gradient shows the direction where there is greatest ratio of change of real value of goods produced in a year (**Y**) depending on person-hours worked in a year (**L**) and real capital change (**K**).

Task (c)

Multivariable functions are used in machine learning to transform high-dimension (multivariable) inputs into a class prediction output.

For this example, we calculate (negative) gradients for each weight parameter in backpropagation and then use them in gradient descent when we would like to calculate in which direction we should move to minimize the loss function (prediction error) the most.

While directional derivates are not usually used in gradient descent, they might be useful if we know a local minimum the loss function might get into when we follow the gradient, and instead we want to direct the gradient descent to "go around" that minimum.

A.2

Heavy Metal (***#diffapplication***, ***#difftheory***)

Task (a)

To find the rate of change of T when moving towards a certain point (directional derivative), we need to first find the vector of movement.

Begin point: (1, 2, 2), end point: (2, 1, 3). Vector of movement: $\vec{v} =< 2-1, 1-2, 3-2 >=< 1, -1, 1 >$

Then, to transform this direction vector into a unit vector, we will divide all its coordinates by its magnitude:

$$||\vec{v}|| = |v| = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}$$
$$\vec{u} = \frac{\vec{v}}{|v|} =< \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} >$$

Since T is inversely proportional to the distance from the center, then we can infer that the formula for the temperature could be:

$$T = \frac{T_0}{1 + \sqrt{x^2 + y^2 + z^2}}$$

where T_0 is the temperature at the origin. 1 is added to the denominator to avoid division by zero when the point is at the origin.

$$\frac{T_0}{1 + \sqrt{1^2 + 2^2 + 2^2}} = 120^\circ$$
$$T_0 = 120 + 120 * 3 = 480^\circ$$

Next step we would like to find all partial derivatives of T to find its gradient:

$$\frac{\partial T}{\partial x} = -\frac{480x}{(x^2 + y^2 + z^2)^{3/2}}$$
$$\frac{\partial T}{\partial y} = -\frac{480y}{(x^2 + y^2 + z^2)^{3/2}}$$
$$\frac{\partial T}{\partial z} = -\frac{480z}{(x^2 + y^2 + z^2)^{3/2}}$$
$$\nabla T =< -\frac{480x}{(x^2 + y^2 + z^2)^{3/2}}, -\frac{480y}{(x^2 + y^2 + z^2)^{3/2}}, -\frac{480z}{(x^2 + y^2 + z^2)^{3/2}} >$$

Now we can compute the directional derivative by "projecting" the gradient onto the direction unit vector:

$$D_{\vec{u}}T(1,2,2) = \nabla T \circ \vec{u} = -480 * \frac{1 * \frac{1}{\sqrt{3}} + 2 * (-\frac{1}{\sqrt{3}}) + 2 * \frac{1}{\sqrt{3}}}{(1+4+4)^{3/2}} = -\frac{480}{\sqrt{3} * 27} = -\frac{160}{9\sqrt{3}}(^{\circ})$$

Task (b)

First, let's find out why the gradient vector is pointing to the origin in this case. As we've discussed before, the gradient vector is:

$$\nabla T =< -\frac{480x}{(x^2+y^2+z^2)^{3/2}}, -\frac{480y}{(x^2+y^2+z^2)^{3/2}}, -\frac{480z}{(x^2+y^2+z^2)^{3/2}} >$$

$$\nabla T = \frac{480}{(x^2+y^2+z^2)^{3/2}} < -x, -y, -z >$$

Recalling that ∇T is a gradient from coordinates (x, y, z), we can easily see that this is a parallel vector with the opposite direction, scaled by some positive value. Since $< x, y, z >$ is the vector lying on the line going through (0, 0, 0) and (x, y, z), $k* < -x, -y, -z >$ would be a vector lying on a parallel line. Since we know this this vector passes through (x, y, z) (since we calculate gradient from that point), it will surely pass (0, 0, 0), meaning that it points "towards" the origin of the metal ball.

When we find the directional derivative from any point in the ball, we project the gradient vector on the direction vector. If we project the gradient vector on the vector with the same direction as itself, then the projected length will be maximized.

Gradient itself tells us about how fast the function is changing in each direction. This also gives us information about which axes to "prioritize" because if the function changes the fastest in direction X then we should move along it. In this case, the gradient would point to the center because the most changes would happen when we reduce the sum of squares the most.

Formula-wise, we know that the angle between two vectors can be defined as:

$$cos\theta = \frac{\vec{u} \circ \vec{v}}{|u||v|}$$

$$|u||v|cos\theta = \vec{u} \circ \vec{v}$$

We can see that the left expression is maximized when $cos\theta = 1$ ($\theta = 0$). This means that the right expression is maximized when the angle between the two vectors is zero, which is when they are parallel to each other.

Part B

B.1

Euro Trip (***#diffapplication***)

Task (a)

In [115]:
```
f(x, y) = y / x
plot3d(f, (x, 0.6, 0.9), (y, 150, 200), mesh=True)
```

Out[115]:

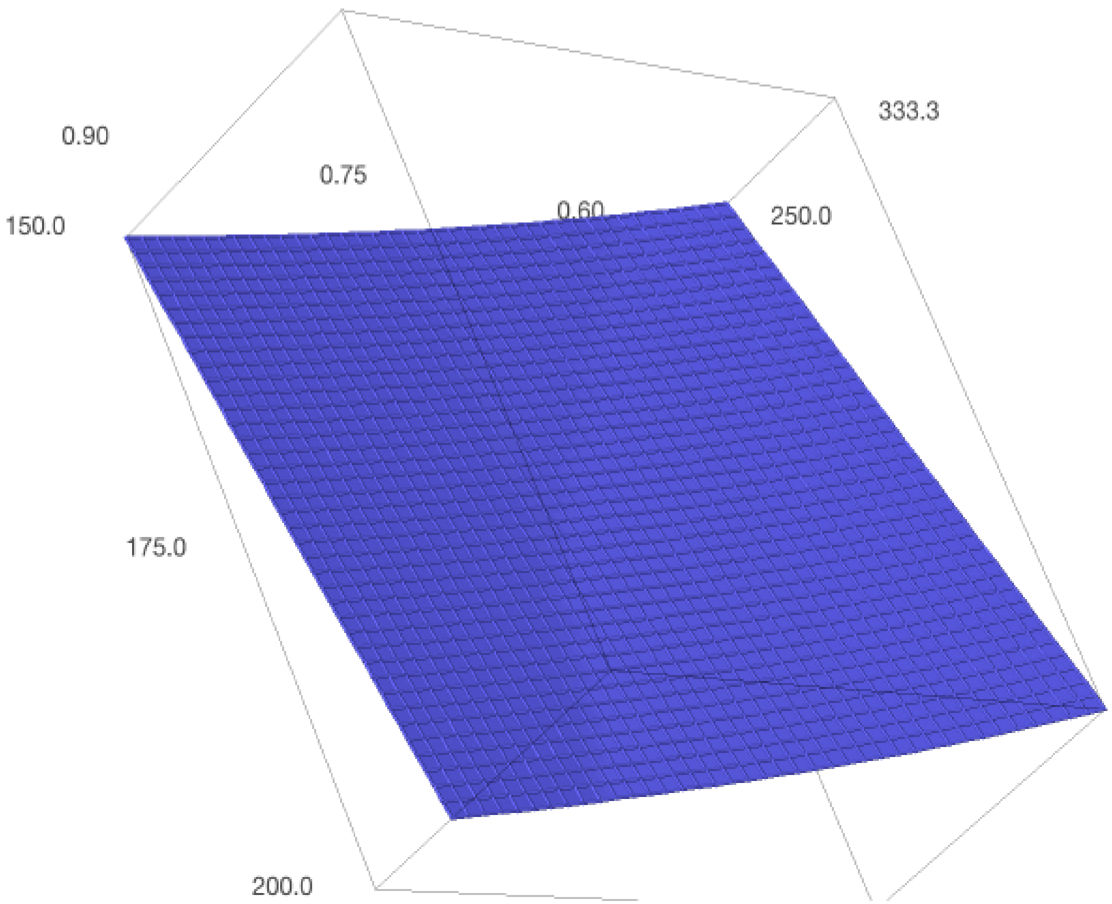

Figure 1. Graph of amount of money in Dollars depending on conversion rate ($\frac{Euro}{Dollar}$) ranging from 0.6 to 0.9 and amount of money exchanged ($Euro$) ranging from 150 to 200.

Task (b)

$$(1)\ \frac{\partial f}{\partial x} = -\frac{y}{x^2}$$
$$(2)\ \frac{\partial f}{\partial y} = \frac{1}{x}$$

Equation (1) indicates the rate of change of amount of dollars w.r.t. the conversion rate.

Equation (2) indicates the rate of change of amount of dollars w.r.t. the amount in Euro we are trying to exchange.

Task (c)

(1) Using related rates

If we introduce the variable t which is the time. Then $\frac{df}{dt}$ would represent the rate of change of the conversion w.r.t. time.

$$f(x, y) = \frac{y}{x}$$
$$\frac{df(x, y)}{dt} = \frac{\frac{dy}{dt}x - \frac{dx}{dt}y}{x^2}, \frac{dy}{dt} = 0$$
$$\frac{df(x, y)}{dt} = -\frac{dx}{dt} * \frac{y}{x^2} = 0.05x\frac{y}{x^2} = 0.05\frac{y}{x}$$

'

This means that when the conversion rate decreases, we will get more Dollars for the Euros we have, specifically, we will get 5% more.

(2) Using partial derivatives

Equation of the plane, tangent at point (x_0, y_0):

$$f_{tan}(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$
$$f_{tan}(x, y) = \frac{y_0}{x_0} - \frac{y_0}{x_0^2}(x - x_0) + \frac{1}{x_0}(y - y_0)$$

Since we get a 5% decrease, we will put in $x = 0.95x_0$:

$$f_{tan}(0.95x_0, y_0) = \frac{y_0}{0.95x_0} + \frac{y_0}{x_0^2}0.05x_0 + 0 = \frac{419}{380}\frac{y_0}{x_0}$$
$$\Delta f = f_{tan}(0.95x_0, y_0) - f(x_0, y_0) \approx 0.10\frac{y_0}{x_0}$$

This is a linear approximation of how the function changes when x decreases by 5% (not so precise, since it should increase by 5%, not 10%).

(3) Using partial derivatives (2)

$$\frac{\partial f(x_0, y_0)}{\partial x} = -\frac{y_0}{x_0^2}$$
$$\partial f(x_0, y_0) = -\partial x\frac{y_0}{x_0^2}$$
$$\partial f(x_0, y_0) = 0.05x\frac{y_0}{x_0^2} = 0.05\frac{y_0}{x_0}$$

Same result as in the first one, but now using the definition of a derivative and differential terms.

Task (d)

In [106]:
```
f(x, y) = y / x
plane(x, y) = f(0.81, 200) - 200/(0.81^2) * (x - 0.81) + 1/0.81 * (y - 200)
a = plot3d(plane, (x, 0.6, 0.9), (y, 140, 220), color="red")
b = plot3d(f, (x, 0.6, 0.9), (y, 140, 220))
intercept = point3d((0.81, 200, f(0.81, 200)), fill="True", size=20, color="green")
a + b + intercept
```

Out[106]:

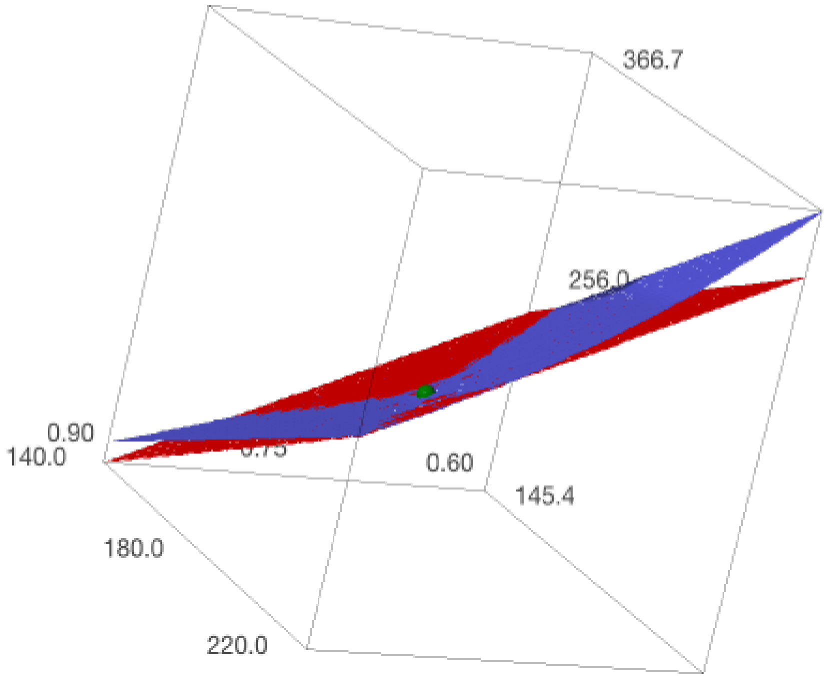

Figure 2. Tangent plane intercepting the conversion function at point (0.81, 200). Using this plane, we then make a linear approximation for the situation when conversion rate is 0.77 $\frac{Euro}{Dollar}$***, and the amount exchanged is 150*** $Euro$***.***

To estimate the combined effects of the discount and the new conversion factor, let's place these values in the tangent plane equation. In this case, we initially knew that conversion rate is 0.81 Euro and product price is 200 Euro and would like to estimate what would be the change if the new conversion rate is 0.77 and product price is 150.

$$f_{tan}(0.77, 150) = \frac{200}{0.81} - \frac{200}{0.81^2}(0.77 - 0.81) + \frac{1}{0.81}(150 - 200) \approx 197.3785$$

Then, the combined effects is the difference between the old price and the new price:

$$\Delta_{est} = |f(0.81, 200) - f_{tan}(0.77, 150)| \approx 49.5351$$

Now, the real difference is:

$$\Delta_{actual} = |f(0.81, 200) - f(0.77, 150)| \approx 52.1084$$

The relative error is:

$$\epsilon = \frac{|\Delta_{est} - \Delta_{actual}|}{\Delta_{actual}} \approx 4.94\%$$

Quite a good estimate, off by only 5%! Most likely what happened here is that the more we go off from the original ratio ($\frac{0.81}{200}$), the less accurate the prediction will be. For example, if we take a look at $(0.6, 220)$ in the graph, the distance is big, because we have decreased the conversion rate, increasing the function, and we increased the converted sum, increasing the function again. Thus, we have diverged from the original ratio, for which we know the exact answer, so then the estimate is worse.

B.2

Mixing Partials (***#difftheory***, ***#limitscontinuity***)

Task (a)

(i)

$$C(x, y) = cos(x^2y^2)$$
$$\frac{\partial C}{\partial x} = -sin(x^2y^2) * 2xy^2$$
$$\frac{\partial^2 C}{\partial y \partial x} = -cos(x^2y^2) * 2yx^2 * 2xy^2 - 4xy * sin(x^2y^2) = -4xy(x^2y^2cos(x^2y^2) + sin(x^2y^2))$$

$$\frac{\partial C}{\partial y} = -sin(x^2y^2) * 2yx^2$$

$$\frac{\partial^2 C}{\partial x \partial y} = -cos(x^2y^2) * 2xy^2 * 2yx^2 - 4yx * sin(x^2y^2) = -4xy(x^2y^2cos(x^2y^2) + sin(x^2y^2))$$

We can observe that

$$\frac{\partial^2 C}{\partial x \partial y} = \frac{\partial^2 C}{\partial y \partial x}$$

(ii)

$$r(\theta, t) = \frac{1}{sin^2(\theta) + e^{-t}}$$

$$\frac{\partial r}{\partial \theta} = -\frac{1}{(sin^2\theta + e^{-t})^2} * 2sin\theta * cos\theta$$

$$\frac{\partial^2 r}{\partial t \partial \theta} = -\frac{1}{(sin^2\theta + e^{-t})^3} * 4sin\theta * cos\theta * e^{-t}$$

$$\frac{\partial r}{\partial t} = -\frac{1}{(sin^2\theta + e^{-t})^2} * e^{-t}$$

$$\frac{\partial^2 r}{\partial t \partial \theta} = -\frac{1}{(sin^2\theta + e^{-t})^3} * 4sin\theta * cos\theta * e^{-t}$$

We can observe that

$$\frac{\partial^2 r}{\partial \theta \partial t} = \frac{\partial^2 r}{\partial t \partial \theta}$$

Checking answers with Sage

In [25]:
```
C(x, y) = cos(x^2*y^2)
C.diff(x).diff(y)
```

Out[25]: `(x, y) |--> -4*x^3*y^3*cos(x^2*y^2) - 4*x*y*sin(x^2*y^2)`

In [26]:
```
r(theta, t) = 1 / (sin(theta)^2 + e^(-t))
r.diff(theta).diff(t)
```

Out[26]:
```
(theta, t) |--> -4*cos(theta)*e^(-t)*sin(theta)/(sin(theta)^2 + e^(-t))^3
```

Task (b)

In [160]:
```
f(x, y) = x * y * (x^2 - y^2) / (x^2 + y^2)
plot3d(f, (x, -10, 10), (y, -10, 10), mesh=True)
```

Out[160]:

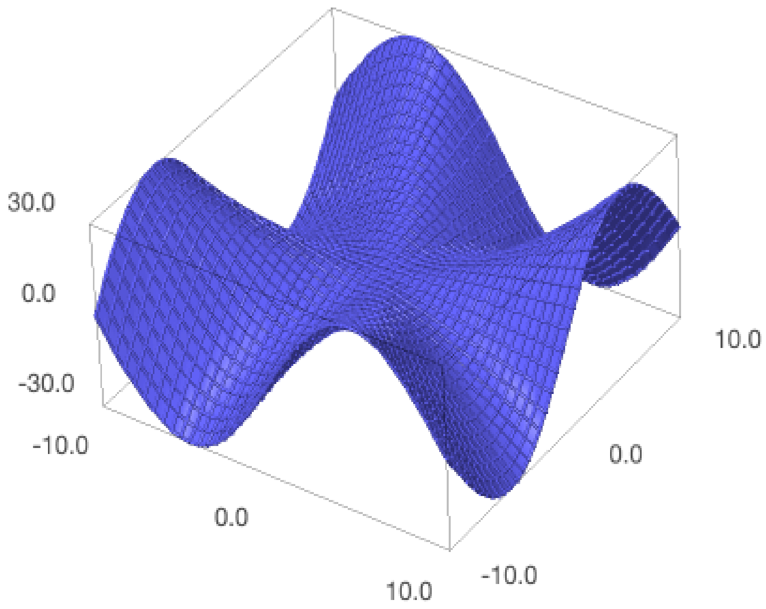

Figure 3. Graph of function $f(x, y) = \dfrac{xy(x^2 - y^2)}{x^2 + y^2}$. This function is continuous at origin (0, 0), since limits from different paths equal to value of the function at that point.

(i)

f is continuous on $\mathbb{R}^2$ because point (0, 0) is defined separately, and there the function is continuous there because $\lim_{(x,y)\to(0,0)} f = 0$ (calculated below). Function f is a composition of elementary functions that are differentiable on their domain, therefore, it is differentiable on its domain, which is $\mathbb{R}^2$.

In [134]:

```
# from multiple paths the function approaches zero
# most likely, the limit at (0, 0) would be 0
print(lim(f(x, 0), x=0, dir='-'))
print(lim(f(x, 0), x=0, dir='+'))
print(lim(f(x, x), x=0))
print(lim(f(x, x/2), x=0))
print(lim(f(0, y), y=0, dir='-'))
print(lim(f(0, y), y=0, dir='+'))
```

```
0
0
0
0
0
0
```

(ii)

$$\frac{\partial f}{\partial x} = \frac{2x^2y}{x^2+y^2} - \frac{2x^2y(x^2-y^2)}{(x^2+y^2)^2} + \frac{y(x^2-y^2)}{x^2+y^2} = \frac{y}{x^2+y^2}(3x^2 - y^2 - \frac{2x^4 - 2x^2y^2}{x^2+y^2})$$

$$\frac{\partial f}{\partial y} = -\frac{2xy^2}{x^2+y^2} - \frac{2xy^2(x^2-y^2)}{(x^2+y^2)^2} + \frac{x(x^2-y^2)}{x^2+y^2} = \frac{x}{x^2+y^2}(x^2 - 3y^2 - \frac{2x^2y^2 - 2y^4}{x^2+y^2})$$

In [135]:

```
# checking
print(f.diff(x))
print(f.diff(y))
```

```
(x, y) |--> 2*x^2*y/(x^2 + y^2) - 2*(x^2 - y^2)*x^2*y/(x^2 + y^2)^2 + (x^2 - y^2)*y/(x^2 + y^2)
(x, y) |--> -2*x*y^2/(x^2 + y^2) - 2*(x^2 - y^2)*x*y^2/(x^2 + y^2)^2 + (x^2 - y^2)*x/(x^2 + y^2)
```

(iii)

$$\frac{\partial f(0,0)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\frac{y(x+\Delta x)((x+\Delta x)^2 - y^2)}{(x+\Delta x)^2 + y^2} - \frac{xy(x^2-y^2)}{x^2+y^2}}{\Delta x} =$$

$$\lim_{\Delta x \to 0} \frac{0-0}{\Delta x} = 0$$

(We fixed y = 0 and placed it in the expression)

$$\frac{\partial f(0,0)}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} = \lim_{\Delta y \to 0} \frac{\frac{x(y+\Delta y)(x^2 - (y+\Delta y)^2)}{x^2 + (y+\Delta y)^2} - \frac{xy(x^2-y^2)}{x^2+y^2}}{\Delta y} =$$

$$\lim_{\Delta y \to 0} \frac{0-0}{\Delta y} = 0$$

(We fixed x = 0 and placed it in the expression)

(iv)

$$\frac{\partial f(x,y)}{\partial x} = \begin{cases} \frac{y}{x^2+y^2}(3x^2 - y^2 - \frac{2x^4 - 2x^2y^2}{x^2+y^2}) & \text{if } (x,y) \neq 0 \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

$$\frac{\partial f(x,y)}{\partial y} = \begin{cases} \frac{x}{x^2+y^2}(x^2 - 3y^2 - \frac{2x^2y^2 - 2y^4}{x^2+y^2}) & \text{if } (x,y) \neq 0 \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

(v)

Using limit definition of derivatives and placing values in the piecewise defined functions in (iv):

$$\frac{\partial^2 f(0,0)}{\partial y \partial x} = lim_{\Delta y \to 0} \frac{\frac{\partial f(0,\Delta y)}{\partial x} - \frac{\partial f(0,0)}{\partial x}}{\Delta y} = lim_{\Delta y \to 0} \frac{\frac{\Delta y * (-\Delta y^2)}{\Delta y^2}}{\Delta y} = lim_{\Delta y \to 0} \frac{-\Delta y^3}{\Delta y^3} = -1$$

$$\frac{\partial^2 f(0,0)}{\partial x \partial y} = lim_{\Delta x \to 0} \frac{\frac{\partial f(\Delta x, 0)}{\partial y} - \frac{\partial f(0,0)}{\partial y}}{\Delta x} = 1$$

(vi)

If we look at the mixed derivates formula (cell below), we have $(x^2 + y^2)$ in the denominator, which means that they might not be continuous at (0, 0). In fact, they are not continuous because the limits from different paths are different. Therefore, we have not met the requirement for Clairaut's theorem that f(x, y) has continuous second partial (mixed) derivate at point (0, 0). Hence, the symmetry does not have to work.

In [154]:

```
g(x, y) = f.diff(x).diff(y)
print "Path x/2 and x, when x -> 0:", lim(g(x/2, x), x = 0)
print "Path x and x, when x -> 0:", lim(g(x, x), x = 0)
```

```
Path x/2 and x, when x -> 0: -171/125
Path x and x, when x -> 0: 0
```

Appendix

#deduction: I employed formal logic in A.2.b and B.2.b by finding the sufficient premises to proof/dispute the provided theorems.
#descriptivestats: I calculated the relative error of linear approximation in A.1.a and B.1.d and interpreted whether the quality of the estimate and the reason why the relative error is as such.
#scienceoflearning: I have engaged in deliberate practice by going through Prof.'s feedback in the previous assignment. In particular, I have focused more on providing clear premises and conclusions step-by-step for my deductive proofs/disputes.