Math 2321 (Multivariable Calculus)

Lecture #8 of 38 $\sim$ February 4, 2021

Tangent Lines and Planes + Linearization

- Geometry of Directional Derivatives + Gradient Vectors
- Tangent Lines and Planes
- Linearization

This material represents §2.2.2 + 2.4.1 from the course notes.
Note that this is slightly out of order. §2.3 will be next week.

Recall, I

Last time, we defined directional derivatives:

Definition

If $\mathbf{v} = \langle v_x, v_y \rangle$ is a unit vector, then the <u>directional derivative of $f(x, y)$ in the direction of $\mathbf{v}$</u> at (x, y), denoted $D_{\mathbf{v}}(f)(x, y)$, is defined to be the limit $D_{\mathbf{v}}(f)(x, y) = \lim_{h \to 0} \dfrac{f(x + h\,v_x,\, y + h\,v_y) - f(x, y)}{h}$, provided that the limit exists.

The directional derivative $D_{\mathbf{v}}(f)(\mathbf{x})$ measures the rate of change of the function f at the point $\mathbf{x}$ in the direction of the unit vector $\mathbf{v}$.

Recall, II

We can compute directional derivatives easily using the gradient:

Definition

The gradient of a function $f(x, y)$, denoted ∇f, is the vector-valued function $\nabla f(x, y) = \langle f_x(x, y),\ f_y(x, y) \rangle$. For a function $g(x, y, z)$, the gradient ∇g is $\nabla g(x, y, z) = \langle g_x(x, y, z),\ g_y(x, y, z),\ g_z(x, y, z) \rangle$.

Theorem (Gradient and Directional Derivatives)

If $\mathbf{v}$ is any unit vector, and f is a function all of whose partial derivatives are continuous, then the directional derivative $D_{\mathbf{v}} f$ satisfies $D_{\mathbf{v}} f = \nabla f \cdot \mathbf{v}$. In words, the directional derivative of f in the direction of $\mathbf{v}$ is the dot product of ∇f with the vector $\mathbf{v}$.

Recall, III

Using the dot product theorem, we deduced the following:

Corollary (Minimum and Maximum Increase)

Suppose f is a differentiable function with gradient $\nabla f \neq \mathbf{0}$ and $\mathbf{v}$ is a unit vector. Then the following hold:

1. *The maximum value of $D_{\mathbf{v}} f$ occurs when $\mathbf{v}$ is a unit vector in the direction of ∇f, and the maximum value is $||\nabla f||$.*
2. *The minimum value of $D_{\mathbf{v}} f$ occurs when $\mathbf{v}$ is a unit vector in the opposite direction of ∇f, and the minimum is $-||\nabla f||$.*
3. *The value of $D_{\mathbf{v}} f$ is zero if and only if $\mathbf{v}$ is orthogonal to the gradient ∇f.*

Tangent Lines and Planes, I

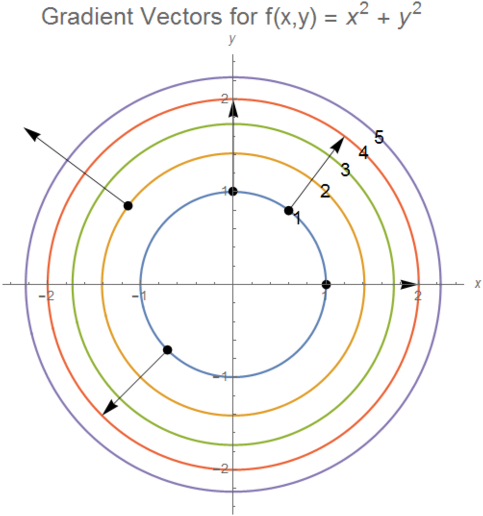
We showed, during our discussion last time, that any direction vector orthogonal to ∇f is necessarily a direction in which the value of the function is not changing.

- In particular, if we imagine traveling along one of the level curves of $f(x, y)$, then by definition the value of f is not changing.
- Therefore, by putting these two observations together, we see that $\nabla f(a, b)$ at that point will be a normal vector to the graph of the level curve at (a, b).
- This is quite readily visible from plots: the gradient vector always runs perpendicular to the level curve.

This means we can use the gradient to find equations of tangent lines to implicit curves described by level sets.

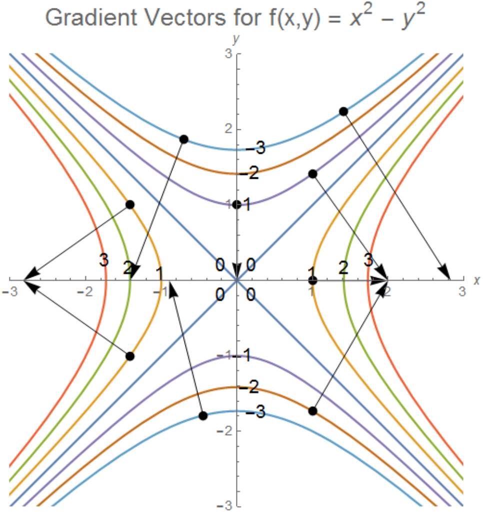
Plots, I

Here are plots of some gradient vectors for $f(x, y) = x^2 + y^2$:



Plots, II

Here are plots of some gradient vectors for $f(x,y) = x^2 - y^2$:



Tangent Lines and Planes, II

Explicitly, suppose that we have an implicit curve of the form $f(x, y) = c$ and we want to find the tangent line to the curve at $(x, y) = (a, b)$.

- Then a vector normal (i.e., orthogonal) to the tangent line is the gradient $\nabla f(a, b)$.
- Then by the same method as we used for describing equations of planes given their normal vector, the tangent line to the curve $f(x, y) = d$ at (a, b) has equation $\nabla f(a, b) \cdot \langle x - a, y - b \rangle = 0$, or, explicitly, $f_x(a, b) \cdot (x - a) + f_y(a, b) \cdot (y - b) = 0$.
- Note that we could also have calculated the slope of this tangent line via (calculus-1-style) implicit differentiation. Both methods, of course, will give the same answer, but I think this procedure is much easier.

Tangent Lines and Planes, III

Example: Find an equation for the tangent line to the curve $x^3 + y^4 = 2$ at the point $(1, 1)$.

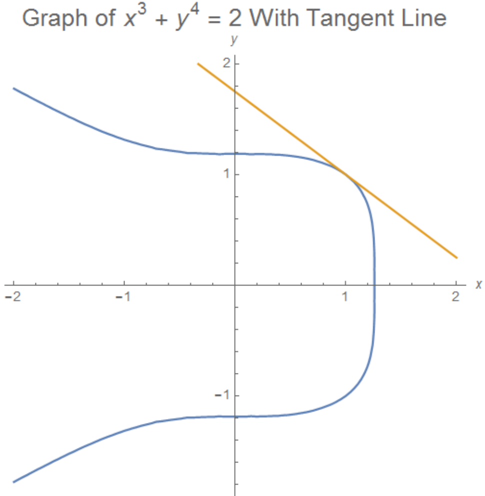
Tangent Lines and Planes, III

Example: Find an equation for the tangent line to the curve $x^3 + y^4 = 2$ at the point $(1, 1)$.

- This curve is the level set $f(x, y) = 2$ for $f(x, y) = x^3 + y^4$, and we have $(a, b) = (1, 1)$.
- We have $f_x = 3x^2$ and $f_y = 4y^3$ so $\nabla f = \langle 3x^2, 4y^3 \rangle$, and $\nabla f(1, 1) = \langle 3, 4 \rangle$.
- Therefore, we get the equation $3(x - 1) + 4(y - 1) = 0$ for the tangent line.

Tangent Lines and Planes, IV

Here is a plot of $x^3 + y^4 = 2$ with its tangent line $3(x-1) + 4(y-1) = 0$ at the point $(1, 1)$:



Tangent Lines and Planes, V

We can use essentially the same procedure to find the equation of the tangent plane to an implicit surface $f(x, y, z) = d$ at a point $(x, y, z) = (a, b, c)$.

Tangent Lines and Planes, V

We can use essentially the same procedure to find the equation of the tangent plane to an implicit surface $f(x, y, z) = d$ at a point $(x, y, z) = (a, b, c)$.

- The point is that the gradient ∇f is orthogonal to the tangent plane, so we can take it as the plane's normal vector.
- Recall that the equation of the plane with normal vector $\mathbf{v} = \langle v_x, v_y, v_z \rangle$ passing through (a, b, c) is $v_x(x - a) + v_y(y - b) + v_z(z - c) = 0$.
- Thus, the tangent plane to the surface $f(x, y, z) = d$ at the point (a, b, c) has equation $\nabla f(a, b, c) \cdot \langle x - a, y - b, z - c \rangle = 0$.
- Explicitly, the tangent plane is $f_x(a, b, c) \cdot (x - a) + f_y(a, b, c) \cdot (y - b) + f_z(a, b, c) \cdot (z - c) = 0$.

Tangent Lines and Planes, VI

Example: Find an equation for the tangent plane to the surface $x^4 + y^4 + z^2 = 3$ at the point $(-1, -1, 1)$.

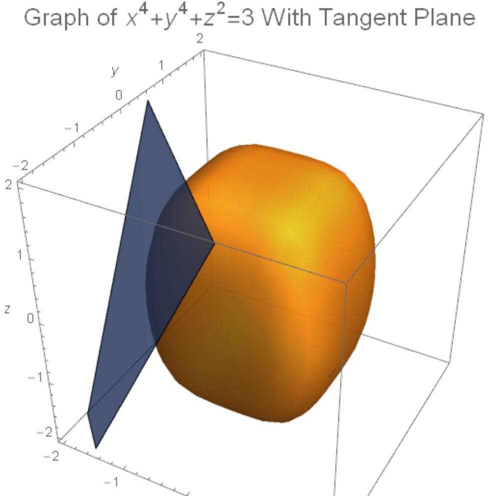
Tangent Lines and Planes, VI

Example: Find an equation for the tangent plane to the surface $x^4 + y^4 + z^2 = 3$ at the point $(-1, -1, 1)$.

- This surface is the level set $f(x, y, z) = 3$ for $f(x, y, z) = x^4 + y^4 + z^2$, and we have $(a, b, c) = (-1, -1, 1)$.
- We have $f_x = 4x^3$, $f_y = 4y^3$, and $f_z = 2z$ so $\nabla f = \left\langle 4x^3, 4y^3, 2z \right\rangle$, and $\nabla f(-1, -1, 1) = \langle -4, -4, 2 \rangle$.
- Therefore, an equation for the tangent plane is $-4(x + 1) - 4(y + 1) + 2(z - 1) = 0$.
- We can rewrite this equation in various ways, such as $-4x - 4y + 2z = 10$ or as $2x + 2y - z = -5$.

Tangent Lines and Planes, VII

Here is a plot of $x^4 + y^4 + z^2 = 3$ with its tangent plane $-4(x+1) - 4(y+1) + 2(z-1) = 0$ at the point $(-1, -1, 1)$:



Tangent Lines and Planes, VIII

Example: Find an equation for the tangent plane to $x^2yz^3 + 2y^3z = 12$ at the point $(x, y, z) = (1, 1, 2)$.

Tangent Lines and Planes, VIII

<u>Example</u>: Find an equation for the tangent plane to $x^2yz^3 + 2y^3z = 12$ at the point $(x, y, z) = (1, 1, 2)$.

- This curve is the level set $f(x, y, z) = 12$ for $f(x, y, z) = x^2yz^3 + 2y^3z$.
- We have $f_x = 2xyz^3$, $f_y = x^2z^3 + 6y^2z$, $f_z = 3x^2yz^2 + 2y^3$, so $\nabla f = \left\langle 2xyz^3, x^2z^3 + 6y^2z, 3x^2yz^2 + 2y^3 \right\rangle$ and $\nabla f(1, 1, 2) = \langle 16, 20, 14 \rangle$.
- Therefore, an equation for the tangent plane is $16(x - 1) + 20(y - 1) + 14(z - 2) = 0$.

Tangent Lines and Planes, IX

Example: Find an equation for the tangent plane to $z = \ln(2x^2 - y^2)$ at the point with $(x, y) = (-1, 1)$.

Tangent Lines and Planes, IX

Example: Find an equation for the tangent plane to $z = \ln(2x^2 - y^2)$ at the point with $(x, y) = (-1, 1)$.

- Note that we must first rearrange the given equation to have the form $f(x, y, z) = d$.
- This curve is the level set $f(x, y, z) = 0$ for $f(x, y, z) = \ln(2x^2 - y^2) - z$.

Tangent Lines and Planes, IX

Example: Find an equation for the tangent plane to $z = \ln(2x^2 - y^2)$ at the point with $(x, y) = (-1, 1)$.

- Note that we must first rearrange the given equation to have the form $f(x, y, z) = d$.
- This curve is the level set $f(x, y, z) = 0$ for $f(x, y, z) = \ln(2x^2 - y^2) - z$. When $(x, y) = (-1, 1)$ we see $z = \ln(2 - 1) = 0$, and so $(a, b, c) = (-1, 1, 0)$.
- We have $f_x = \dfrac{4x}{2x^2 - y^2}$, $f_y = \dfrac{-2y}{2x^2 - y^2}$, and $f_z = -1$ so $\nabla f = \left\langle \dfrac{4x}{2x^2 - y^2}, \dfrac{-2y}{2x^2 - y^2}, -1 \right\rangle$ and $\nabla f(-1, 1, 0) = \langle -4, -2, -1 \rangle$.
- Therefore, an equation for the tangent plane is $-4(x + 1) - 2(y - 1) - (z) = 0$, or equivalently $-4x - 2y - z = 2$.

Linearization, I

Just like with functions of a single variable, we have a notion of a "best linear approximation" and a "best polynomial approximation" to a function of several variables.

- These ideas are often needed in applications in the sciences, engineering, and applied mathematics, when it is frequently necessary to analyze a complicated system with a series of approximations.
- In particular, taking a "first-order approximation" to solving a problem means to write down the linearized version (and then solve it).
- Pleasantly, the general notion of linearization is quite closely tied to the gradient and tangent lines/planes.

Linearization, II

The key insight is that the tangent plane to the graph of a function is "the best linear approximation" to that function near the point of tangency. Here is a justification of this idea:

- Suppose we want to compute the change in a function $f(x, y)$ as we move from (a, b) to a nearby point $(a + \Delta x,\ b + \Delta y)$.
- A slight modification of the definition of the directional derivative says that, for $\mathbf{v} = \langle \Delta x,\ \Delta y \rangle$, we have $||\mathbf{v}||\ D_{\mathbf{v}} f(a, b) = \lim_{h \to 0} \dfrac{f(a + h\Delta x,\ b + h\Delta y) - f(a, b)}{h}$.
- When Δx and Δy are small, then the difference quotient should be close to the limit value.
- From the properties of the gradient, we know $D_{\mathbf{v}} f(a, b) = \nabla f(a, b) \cdot \dfrac{\mathbf{v}}{||\mathbf{v}||} = \dfrac{1}{||\mathbf{v}||} \left[f_x(a, b)\Delta x + f_y(a, b)\Delta y\right]$.

Linearization, II

The key insight is that the tangent plane to the graph of a function is "the best linear approximation" to that function near the point of tangency. Here is a justification of this idea:

- Expanding out the calculation on the previous slide yields $f(a+\Delta x, b+\Delta y) \approx f(a,b) + f_x(a,b)\cdot\Delta x + f_y(a,b)\cdot\Delta y$.
- If we write $x = a + \Delta x$ and $y = a + \Delta y$, we see that $f(x,y)$ is approximately equal to the linear function $L(x,y) = f(a,b) + f_x(a,b)\cdot(x-a) + f_y(a,b)\cdot(y-b)$ when $x-a$ and $y-b$ are small.

Linearization, III

We can summarize this discussion as follows:

Definition

If $f(x, y)$ is a differentiable function of two variables, its *linearization* *at the point $(x, y) = (a, b)$ is the linear function $L(x, y) = f(a, b) + f_x(a, b) \cdot (x - a) + f_y(a, b) \cdot (y - b)$. The linearization is the best linear approximation to $f(x, y)$ near (a, b).*

We also remark that this linearization is the same as the approximation given by the tangent plane, since the tangent plane to $z = f(x, y)$ at (a, b) has equation $z = f(a, b) + f_x(a, b) \cdot (x - a) + f_y(a, b) \cdot (y - b)$.

- In other words, the tangent plane to the graph of $z = f(x, y)$ at $(x, y) = (a, b)$ gives a good approximation to the function $f(x, y)$ near the point of tangency.

Linearization, IV

Example: Let $f(x, y) = e^{x+y}$.

1. Find the linearization of f at $(x, y) = (0, 0)$.

Linearization, IV

Example: Let $f(x, y) = e^{x+y}$.

1. Find the linearization of f at $(x, y) = (0, 0)$.

- In this case, we have $(a, b) = (0, 0)$, and we calculate $f_x = e^{x+y}$ and $f_y = e^{x+y}$.
- Therefore, $L(x, y) = f(0, 0) + f_x(0, 0) \cdot (x - 0) + f_y(0, 0) \cdot (y - 0) = 1 + x + y$.

Linearization, IV

Example: Let $f(x, y) = e^{x+y}$.

1. Find the linearization of f at $(x, y) = (0, 0)$.

- In this case, we have $(a, b) = (0, 0)$, and we calculate $f_x = e^{x+y}$ and $f_y = e^{x+y}$.
- Therefore, $L(x, y) = f(0, 0) + f_x(0, 0) \cdot (x - 0) + f_y(0, 0) \cdot (y - 0) = 1 + x + y$.

2. Use the linearization of f to estimate $f(0.1, 0.1)$.

Linearization, IV

Example: Let $f(x, y) = e^{x+y}$.

1. Find the linearization of f at $(x, y) = (0, 0)$.

- In this case, we have $(a, b) = (0, 0)$, and we calculate $f_x = e^{x+y}$ and $f_y = e^{x+y}$.
- Therefore, $L(x, y) = f(0, 0) + f_x(0, 0) \cdot (x - 0) + f_y(0, 0) \cdot (y - 0) = 1 + x + y$.

2. Use the linearization of f to estimate $f(0.1, 0.1)$.

- The point is that $L(x, y)$ should be a good estimate for (x, y) when (x, y) is close to $(0, 0)$.
- The approximate value of $f(0.1, 0.1)$ is thus $L(0.1, 0.1) = 1.2$.
- The actual value is $e^{0.2} \approx 1.2214$, which is reasonably close.

Linearization, VI

Example: Let $f(x, y) = \sqrt[3]{x + 3y^2}$.

1. Find the linearization of f at $(x, y) = (5, 1)$.

Linearization, VI

Example: Let $f(x,y) = \sqrt[3]{x+3y^2}$.

1. Find the linearization of f at $(x,y) = (5,1)$.

- In this case, we have $(a,b) = (5,1)$, and we calculate $f_x = \frac{1}{3}(x+3y^2)^{-2/3}$ and $f_y = \frac{1}{3}(x+3y^2)^{-2/3}(6y)$.
- Therefore,

$$\begin{aligned} L(x,y) &= f(5,1) + f_x(5,1)\cdot(x-5) + f_y(5,1)\cdot(y-1) \\ &= 2 + \frac{1}{12}(x-5) + \frac{1}{2}(y-1). \end{aligned}$$

Linearization, VI

Example: Let $f(x, y) = \sqrt[3]{x + 3y^2}$.

1. Find the linearization of f at $(x, y) = (5, 1)$.

- In this case, we have $(a, b) = (5, 1)$, and we calculate $f_x = \frac{1}{3}(x + 3y^2)^{-2/3}$ and $f_y = \frac{1}{3}(x + 3y^2)^{-2/3}(6y)$.
- Therefore,
$$\begin{aligned} L(x, y) &= f(5, 1) + f_x(5, 1) \cdot (x - 5) + f_y(5, 1) \cdot (y - 1) \\ &= 2 + \frac{1}{12}(x - 5) + \frac{1}{2}(y - 1). \end{aligned}$$

2. Use the linearization of f to estimate $f(5.3, 0.9)$.

Linearization, VI

<u>Example</u>: Let $f(x, y) = \sqrt[3]{x + 3y^2}$.

1. Find the linearization of f at $(x, y) = (5, 1)$.

- In this case, we have $(a, b) = (5, 1)$, and we calculate $f_x = \frac{1}{3}(x + 3y^2)^{-2/3}$ and $f_y = \frac{1}{3}(x + 3y^2)^{-2/3}(6y)$.
- Therefore,
$$\begin{aligned} L(x, y) &= f(5, 1) + f_x(5, 1) \cdot (x - 5) + f_y(5, 1) \cdot (y - 1) \\ &= 2 + \tfrac{1}{12}(x - 5) + \tfrac{1}{2}(y - 1). \end{aligned}$$

2. Use the linearization of f to estimate $f(5.3, 0.9)$.

- The approximate value is
$$L(5.3, 0.9) = 2 + \frac{1}{12}(0.3) + \frac{1}{2}(-0.1) = 1.975.$$
- The actual value is $\sqrt[3]{7.73} \approx 1.97724$, reasonably close.

Linearization, VII

We can, of course, linearize functions of more variables as well:

Definition

If $f(x, y, z)$ is a differentiable function of three variables, its linearization at the point $(x, y) = (a, b, c)$ is the linear function $L(x, y, z) = f(a, b, c) + f_x(a, b, c) \cdot (x - a) + f_y(a, b, c) \cdot (y - b) + f_z(a, b, c) \cdot (z - c)$. *The linearization is the best linear approximation to $f(x, y, z)$ near (a, b, c).*

The point here is that we simply gain the corresponding term for the extra variable z.

Linearization, VIII

Example: Let $f(x, y, z) = x^2 y^3 z^4$.

1. Find the linearization of f at $(x, y, z) = P = (1, 1, 1)$.

Linearization, VIII

Example: Let $f(x, y, z) = x^2y^3z^4$.

1. Find the linearization of f at $(x, y, z) = P = (1, 1, 1)$.

- We have $(a, b, c) = (1, 1, 1)$, and we calculate $f_x = 2xy^3z^4$, $f_y = 3x^2y^2z^4$, and $f_z = 4x^2y^3z^4$.
- Then

$$\begin{aligned} L(x, y, z) &= f(P) + f_x(P) \cdot (x - 1) + f_y(P) \cdot (y - 1) + f_z(P) \cdot (z - 1) \\ &= 1 + 2(x - 1) + 3(y - 1) + 4(z - 1). \end{aligned}$$

Linearization, VIII

<u>Example</u>: Let $f(x, y, z) = x^2y^3z^4$.

1. Find the linearization of f at $(x, y, z) = P = (1, 1, 1)$.

- We have $(a, b, c) = (1, 1, 1)$, and we calculate $f_x = 2xy^3z^4$, $f_y = 3x^2y^2z^4$, and $f_z = 4x^2y^3z^4$.
- Then

$$\begin{aligned} L(x, y, z) &= f(P) + f_x(P) \cdot (x-1) + f_y(P) \cdot (y-1) + f_z(P) \cdot (z-1) \\ &= 1 + 2(x-1) + 3(y-1) + 4(z-1). \end{aligned}$$

2. Use the linearization of f to estimate $f(1.2, 1.1, 0.9)$.

Linearization, VIII

Example: Let $f(x, y, z) = x^2 y^3 z^4$.

1. Find the linearization of f at $(x, y, z) = P = (1, 1, 1)$.
   - We have $(a, b, c) = (1, 1, 1)$, and we calculate $f_x = 2xy^3z^4$, $f_y = 3x^2y^2z^4$, and $f_z = 4x^2y^3z^4$.
   - Then
   $$\begin{aligned} L(x, y, z) &= f(P) + f_x(P)\cdot(x-1) + f_y(P)\cdot(y-1) + f_z(P)\cdot(z-1) \\ &= 1 + 2(x-1) + 3(y-1) + 4(z-1). \end{aligned}$$
2. Use the linearization of f to estimate $f(1.2, 1.1, 0.9)$.
   - The approximate value of $f(1.2, 1.1, 0.9)$ is then $L(1.2, 1.1, 0.9) = 1.3$.
   - The actual value is ≈ 1.2575: again, fairly close.

Linearization, IX

Example: Use a linearization to approximate the change in $f(x, y, z) = e^{x+y}(y + z)^2$ in moving from $(-1, 1, 1)$ to $(-0.9, 0.9, 1.2)$.

Linearization, IX

Example: Use a linearization to approximate the change in $f(x, y, z) = e^{x+y}(y+z)^2$ in moving from $(-1, 1, 1)$ to $(-0.9, 0.9, 1.2)$.

- First we compute the linearization: we have $f_x = e^{x+y}(y+z)^2$, $f_y = e^{x+y}(y+z)^2 + 2e^{x+y}(y+z)$, and $f_z = 2e^{x+y}(y+z)$, so $\nabla f(-1, 1, 1) = \langle 4, 8, 4 \rangle$, and then the linearization is $L(x, y, z) = 4 + 4(x+1) + 8(y-1) + 4(z-1)$.

Linearization, IX

Example: Use a linearization to approximate the change in $f(x, y, z) = e^{x+y}(y+z)^2$ in moving from $(-1, 1, 1)$ to $(-0.9, 0.9, 1.2)$.

- First we compute the linearization: we have $f_x = e^{x+y}(y+z)^2$, $f_y = e^{x+y}(y+z)^2 + 2e^{x+y}(y+z)$, and $f_z = 2e^{x+y}(y+z)$, so $\nabla f(-1, 1, 1) = \langle 4, 8, 4 \rangle$, and then the linearization is $L(x, y, z) = 4 + 4(x+1) + 8(y-1) + 4(z-1)$.
- We see that the approximate change is then $L(-0.9, 0.9, 1.2) - f(-1, 1, 1) = 4.4 - 4 = 0.4$.
- We could also have estimated this change using a directional derivative: the result is $\nabla f(-1, 1, 1) \cdot \langle 0.1, -0.1, 0.2 \rangle = 0.4$.
- This estimate is exactly the same as the one arising from the linearization; this should not be surprising, since the two calculations are ultimately the same.

Taylor Series, I [FOR FUN ONLY]

In approximating a function by its best linear approximation, we might like to be able to bound how far off our approximations are: after all, an approximation is not very useful if we do not know how good it is!

- One can give an upper bound on the error using Taylor series and the multivariable version of Taylor's Remainder Theorem.
- The general motivation for multivariable Taylor series is that there is no reason only to consider linear approximations, aside from the fact that linear functions are easiest: we could just as well ask about how to approximate $f(x, y)$ with a higher-degree polynomial in x and y.
- Like in the one-variable case, there is a very natural notion of Taylor series that arises by summing over higher derivatives of the function.

Taylor Series, II [FOR FUN ONLY]

Here's the definition for a function of two variables:

Definition

If $f(x,y)$ is a function all of whose nth-order partial derivatives at $(x,y)=(a,b)$, the Taylor series for $f(x)$ at $(x,y)=(a,b)$ is

$$T(x,y)=\sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(x-a)^k(y-b)^{n-k}}{k!(n-k)!}\frac{\partial^n f}{(\partial x)^k(\partial y)^{n-k}}(a,b).$$

The degree-d Taylor polynomial is the sum for $n+k\le d$.

- At $(a,b)=(0,0)$ the series is $T(0,0)=f(0,0)+[xf_x+yf_y]+\frac{1}{2!}\left[x^2f_{xx}+2xy\,f_{xy}+y^2f_{yy}\right]+\frac{1}{3!}\left[x^3f_{xxx}+3x^2y\,f_{xxy}+3xy^2f_{xyy}+y^3f_{yyy}\right]+\cdots.$

You can see this is rather complicated even in the simplest case!

Taylor Series, III [FOR FUN ONLY]

We have a multivariable version of Taylor's Remainder Theorem, which provides an upper bound on the error from an approximation of a function by its Taylor polynomial:

Theorem (Taylor's Remainder Theorem, multivariable)

If $f(x, y)$ has continuous partial derivatives up to order $d+1$ near (a, b), and if $T_d(x, y)$ is the degree-d Taylor polynomial for $f(x, y)$ at (a, b), then for any point (x, y), we have
$|T_k(x, y) - f(x, y)| \le M \cdot \dfrac{(|x-a| + |y-b|)^{k+1}}{(k+1)!}$, *where M is a constant such that $|f_\star| \le M$ for every $(d+1)$-order partial derivative $f_\star$ on the segment joining (a, b) to (x, y).*

The proof follows by applying the one-variable version to f along the line segment joining (a, b) to (x, y).

Taylor Series, IV [FOR FUN ONLY]

The utility of the theorem is not so much about approximating particular values of the function f, but about giving a uniform error bound for a polynomial approximation to f on a particular region.

- For example, consider $f(x, y) = e^x \cos(y)$ near $(0, 0)$.
- The best quadratic approximation to f is $T_2(x, y) = 1 + x + \frac{x^2}{2} - \frac{y^2}{2}$.
- One can also check that for $|x|, |y| \leq 0.1$, the value $M = e^{0.1}$ is an upper bound on all the partial derivatives.
- Then Taylor's theorem yields the error estimate $|T_2(x, y) - f(x, y)| \leq e^{0.1} \frac{(0.2)^3}{3!} < 0.0015$ when $|x|, |y| \leq 0.1$.
- The utility of this result is that we get a uniform approximation to the complicated function f by the simple function T_2 with an explicit bound on the error.

Summary

We discussed the relationship between tangent lines/planes and the gradient vector.

We discussed linearization and its relationship to the gradient.

Next lecture: The chain rule and implicit differentiation.