CHAPTER 13  PARTIAL DERIVATIVES

13.1 Surfaces and Level Curves

The graph of \( z = f(x, y) \) is a surface in \( xyz \) space. When \( f \) is a linear function, the surface is flat (a plane). When \( f = x^2 + y^2 \) the surface is curved (a parabola is revolved to make a bowl). When \( f = \sqrt{x^2 + y^2} \) the surface is pointed (a cone resting on the origin). These three examples carry you a long way.

To visualize a surface we cut through it by planes. Often the cutting planes are horizontal, with the simple equation \( z = c \) (a constant). This plane meets the surface in a **level curve**, and the equation of that curve is \( c = f(x, y) \). The cutting is up at all different heights \( c \), but we move all the level curves down to the \( xy \) plane.

For the bowl \( z = x^2 + y^2 \) the level curves are \( c = x^2 + y^2 \) (circles). For the cone \( z = \sqrt{x^2 + y^2} \) the level curves are \( c = \sqrt{x^2 + y^2} \) (again circles - just square both sides). For the plane \( z = x + y \) the level curves are straight lines \( c = x + y \) (parallel to each other as \( c \) changes).

The collection of level curves in the \( xy \) plane is a **contour map**. If you are climbing on the surface, the map tells you two important things:

1. Which way is up: Perpendicular to the level curve is the steepest direction.
2. How steep the surface is: Divide the change in \( c \) by the distance between level curves.

A climbing map shows the curves at equal steps of \( c \). The mountain is steeper when the level curves are closer.

1. Describe the level curves for the saddle surface \( z = xy \).
   - The curve \( xy = 1 \) is a hyperbola. One branch is in the first quadrant through \((1,1)\). The other branch is in the third quadrant through \((-1,-1)\). At these points the saddle surface has \( z = 1 \).
   - The curve \( xy = -1 \) is also a hyperbola. Its two pieces go through \((1,-1)\) and \((-1,1)\). At these points the surface has \( z = xy = -1 \) and it is below the plane \( z = 0 \).

2. How does a maximum of \( f(x, y) \) show up on the contour map of level curves?
   - Think about the top point of the surface. The highest cutting plane just touches that top point. The level curve is only a point! When the plane moves lower, it cuts out a curve that goes around the top point. So the contour map shows "near-circles" closing in on a single maximum point. A minimum looks just the same, but the \( c \)'s decrease as the contour lines close in.

**Read-throughs and selected even-numbered solutions:**

The graph of \( z = f(x, y) \) is a surface in three-dimensional space. The level curve \( f(x, y) = 7 \) lies down in the base plane. Above this level curve are all points at height \( 7 \) in the surface. The plane \( z = 7 \) cuts through the surface at those points. The level curves \( f(x, y) = c \) are drawn in the \( xy \) plane and labeled by \( c \). The family of labeled curves is a contour map.

For \( z = f(x, y) = x^2 - y^2 \), the equation for a level curve is \( x^2 - y^2 = c \). This curve is a hyperbola. For \( z = x - y \) the curves are straight lines. **Level curves never cross because \( f(x,y) \) cannot equal two numbers \( c \) and \( c' \).** They crowd together when the surface is steep. The curves tighten to a point when \( f \) reaches a maximum or minimum. The steepest direction on a mountain is perpendicular to the level curve.

6 \((x + y)^2 = 0\) gives the line \( y = -x; (x + y)^2 = 1 \) gives the pair of lines \( x + y = 1 \) and \( x + y = -1 \); similarly \( x + y = \sqrt{2} \) and \( x + y = -\sqrt{2} \); no level curve \((x + y)^2 = -4\).
13.2 Partial Derivatives (page 479)

I am sure you are good at taking partial derivatives. They are like ordinary derivatives, when you close your eyes to the other variables. As the text says, "Do not treat y as zero! Treat it as a constant." Just pretend that $y = 5$. That applies to $\frac{\partial}{\partial x} e^{xy} = ye^{xy}$ and $\frac{\partial}{\partial y} (x^2 + xy^2) = 2x + y^2$.

Remember that $\frac{\partial f}{\partial x}$ is also written $f_x$. The y-derivative of this function is $\frac{\partial f}{\partial y}$ or $f_y$. A major point is that $f_{xy} = f_{yx}$. The y-derivative of $\frac{\partial f}{\partial y}$ equals the x-derivative of $\frac{\partial f}{\partial y}$. Take $f = x^2 + xy^2$ with $\frac{\partial f}{\partial y} = 2xy$:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} (2x + y^2) = 2y \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} (2xy) = 2y.$$  

Problem 43 proves this rule $f_{xy} = f_{yx}$, assuming that both functions are continuous. Here is another example:

1. The partial derivatives of $f(x,y) = e^{xy}$ are $f_x = ye^{xy}$ and $f_y = xe^{xy}$. Find $f_{xx}$, $f_{xy}$, $f_{yx}$, and $f_{yy}$.
   - $f_{xx}$ is $\frac{\partial^2 f}{\partial x^2}$ or $\frac{\partial}{\partial x} (\frac{\partial f}{\partial x})$. This is $\frac{\partial}{\partial x} (ye^{xy}) = y^2 e^{xy}$. Similarly $f_{yy}$ is $\frac{\partial}{\partial y} (xe^{xy}) = x^2 e^{xy}$. The mixed derivatives are equal as usual:
     $$\frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial}{\partial y} (ye^{xy}) = y(xe^{xy}) + 1(e^{xy}) \text{ by the product rule}$$
     $$\frac{\partial}{\partial x} \frac{\partial f}{\partial y} = \frac{\partial}{\partial x} (xe^{xy}) = x(ye^{xy}) + 1(e^{xy}) \text{ by the product rule}$$
     
     You must notice that it is $\partial^2 f$ above and $\partial x^2$ below. We divide $\Delta(\Delta f)$ by $(\Delta x)^2$.

2. What does that mean? How is $\Delta(\Delta f)$ different from $(\Delta f)^2$?
   - Start with $f(x)$. The forward difference $\Delta f$ is $f(x + \Delta x) - f(x)$. This is a function of $x$. So we can take its forward difference:
     $$\Delta(\Delta f) = \Delta f(x + \Delta x) - \Delta f(x) = [f(x + 2\Delta x) - f(x + \Delta x)] - [f(x + \Delta x) - f(x)]$$
     
     This is totally different from $(\Delta f)^2 = [f(x + \Delta x) - f(x)]^2$. In the limit $\frac{\partial^2 f}{\partial x^2}$ is totally different from $(\frac{\partial f}{\partial x})^2$.

3. Which third derivatives are equal to $f_{xxy}$? This is $\frac{\partial}{\partial y} (f_{xx})$ or $\frac{\partial^3 f}{\partial y \partial x^2}$.
   - We are taking one y-derivative and two x-derivatives. The order does not matter (for a smooth function). Therefore $f_{xxy} = f_{xyx} = f_{yxx}$.
Notice Problems 45–52 about limits and continuity for functions $f(x, y)$. This two-variable case is more subtle than limits and continuity of $f(x)$. In a course on mathematical analysis this topic would be expanded. In a calculus course I believe in completing the definitions and applying them.

More important in practice are partial differential equations like $\frac{\partial f}{\partial t} = \frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial f}{\partial t} = \frac{\partial^2 f}{\partial x^2}$. Those are the one-way wave equation and the two-way wave equation and the heat equation. Problem 42 says that if $\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2}$ then automatically $\frac{\partial^3 f}{\partial t^3} = \frac{\partial^3 f}{\partial x^3}$. A one-way wave is a special case of a two-way wave.

4. Solve Problem 42. Then find $f(x, t)$ that satisfies the 2-way equation but not the 1-way equation.

- Suppose a particular function satisfies $f_t = f_x$. Take $t$-derivatives to get $f_{tt} = f_{xx}$. Take $x$-derivatives to get $f_{tx} = f_{xx}$. The mixed derivatives agree for any smooth function: $f_{tt} = f_{xx}$. Therefore $f_{tt} = f_{xx}$. Example of a 1-way wave: $f = (x + t)^2$. The function $f = (x + t)^2$ does not satisfy the 1-way equation, because $f_x = 2(x + t)$ and $f_t = -2(x + t)$. It satisfies the other-way wave equation $f_t = -f_x$ with a minus sign. But this is enough for the 2-way equation because $f_{xx} = 2$ and $f_{tt} = 2$.

In general $F(x + t)$ solves the one-way equation, $G(x - t)$ solves the other-way equation, and their sum $F + G$ solves the two-way equation.

Read-throughs and selected even-numbered solutions:

The partial derivative $\partial f/\partial y$ comes from fixing $x$ and moving $y$. It is the limit of $(f(x, y + \Delta y) - f(x, y))/\Delta y$. If $f = e^{2x} \sin y$ then $\partial f/\partial x = 2e^{2x} \sin y$ and $\partial f/\partial y = e^{2x} \cos y$. If $f = (x^2 + y^2)^{1/2}$ then $f_x = x/(x^2 + y^2)^{1/2}$ and $f_y = y/(x^2 + y^2)^{1/2}$. At $(x_0, y_0)$ the partial derivative $f_x$ is the ordinary derivative of the partial function $f(x, y)$. Similarly $f_y$ comes from $f(x_0, y)$. Those functions are cut out by vertical planes $x = x_0$ and $y = y_0$, while the level curves are cut out by horizontal planes.

The four second derivatives are $f_{xx}, f_{xy}, f_{yx}, f_{yy}$. For $f = xy$ they are 0, 1, 1, 0. For $f = \cos 2x \cos 3y$ they are $-4 \cos 2x \cos y, 6 \sin 2x \sin 3y, -9 \cos 2x \cos 3y$. In those examples the derivatives $f_{xy}$ and $f_{yx}$ are the same. That is always true when the second derivatives are continuous. At the origin, $\cos 2x \cos 3y$ is curving down in the $x$ and $y$ directions, while $xy$ goes up in the $45^\circ$ direction and down in the $-45^\circ$ direction.

8 $\frac{\partial f}{\partial x} = \frac{1}{x + 2y}, \quad \frac{\partial f}{\partial y} = \frac{2}{x + 2y}$
18 $f_{xx} = n(n - 1)(x + y)^{n-2} = f_{xy} = f_{yx} = f_{yy}$
20 $f_{xx} = \frac{d}{(x+y)^2}, \quad f_{yy} = \frac{-1}{(x+y)^2}, \quad f_{yy} = \frac{\partial^2 f}{(x+y)^2} = -\frac{2}{(x+y)^2}$ Note $f_{xx} + f_{yy} = 0$.
28 $\frac{\partial f}{\partial x} = -u(x)$ and $\frac{\partial f}{\partial y} = v(y)$.
36 $f_x = \frac{1}{\sqrt{t}} (-\frac{2x}{4t}) e^{-x^2/4t}$. Then $f_{xx} = f_t = \frac{1}{\sqrt{t}} e^{-x^2/4t} + \frac{x^2}{4t^2} e^{-x^2/4t}$.
38 $e^{-m_1 t - n_1 x} \sin m_2 y \cos n_2 y$ solves $f_t = f_{xx} + f_{yy}$. Also $f = \frac{1}{t} e^{-(x^2+y^2)/4t}$ has $f_t = f_{xx} + f_{yy} = \left(-\frac{1}{4t} + \frac{x^2+y^2}{4t^2}\right) e^{-(x^2+y^2)/4t}$.
50 Along $y = mz$ the function is $\frac{ms^2}{x^4 + m^2 x^2 z^2} \to 0$ (the ratio is near $\frac{ms^2}{m^2 z^2}$ for small $z$). But on the parabola $y = x^2$ the function is $\frac{s^2}{2x^2} = \frac{1}{2}$. So this function $f(x, y)$ has no limit: not continuous at $(0,0)$. 

180
13.3 Tangent Planes and Linear Approximations  (page 488)

A smooth curve has tangent lines. The equation of the line uses the derivative (the slope). A smooth surface has tangent planes. The equation of the plane uses two partial derivatives \( f_x \) and \( f_y \) (two slopes). Compare

line \( y - f(a) = f'(a)(x - a) \) with \( z - f(a, b) = f_x(a, b)(x - a) + f_y(a, b)(y - b) \) plane.

These are linear equations. On the left is \( y = mx + \text{constant} \). On the right is \( z = Mx + Ny + \text{constant} \). Linear equations give lines in the \( xy \) plane, and they give planes in \( xyz \) space. The nice thing is that the first slope \( M = \frac{\partial f}{\partial x} \) stays completely separate from the second slope \( N = \frac{\partial f}{\partial y} \).

I will follow up that last sentence. Suppose we change \( a \) by \( Ax \) and \( b \) by \( Ay \). The basepoint is \( (a, b) \) and the movement is to \( (a + Ax, b + Ay) \). Knowing the function \( f \) and its derivatives at the basepoint, we can predict the function (linear approximation) at the nearby point. In one variable we follow the tangent line to \( f(a) + f'(a)Ax \).

In two variables we follow the tangent plane to the nearby point:

We add on two linear corrections, in the \( x \) and \( y \) directions. Often these formulas are written with \( x \) instead of \( a \) and \( y \) instead of \( b \). The movement is from \( f(x, y) \) to \( f(x + \Delta x, y + \Delta y) \). The change is \( \Delta z \) \( f_x \) + \( \Delta y \) \( f_y \).

1. Estimate the change in \( f(x, y) = x^3y^4 \) when you move from \((1,1)\) to \((1 + \Delta x, 1 + \Delta y)\).
   - The \( x \)-derivative is \( f_x = 3x^2y^4 = 3 \) at the basepoint \((1,1)\). The \( y \)-derivative is \( f_y = 4x^3y^3 = 4 \) at the basepoint. The change \( \Delta f \) is approximately \( f_x \Delta x + f_y \Delta y \). This is \( 3\Delta x + 4\Delta y \):

\[
   f(x, y) = (1 + \Delta x)^3(1 + \Delta y)^4 \approx 1 + 3\Delta x + 4\Delta y.
\]

On the left, the high powers \((\Delta x)^3(\Delta y)^4\) would multiply. But the lowest powers \( \Delta x \) and \( \Delta y \) just add. You can see that if you write out \((1 + \Delta x)^3\) and \((1 + \Delta y)^4\) and start multiplying:

\[
   (1 + 3\Delta x + 3(\Delta x)^2 + (\Delta x)^3)(1 + 4\Delta y + \cdots) = 1 + 3\Delta x + 4\Delta y + \text{higher terms}.
\]

These higher terms come into the complete Taylor series. The constant and linear terms are the start of that series. They give the linear approximation.

2. Find the equation of the tangent plane to the surface \( z = x^3y^4 \) at \((x, y) = (1, 1)\).
   - The plane is \( z - 1 = 3(x - 1) + 4(y - 1) \). If \( x - 1 \) is \( \Delta x \) and \( y - 1 \) is \( \Delta y \), this is \( z = 1 + 3\Delta x + 4\Delta y \). Same as Question 1. The tangent plane gives the linear approximation!

Some surfaces do not have “explicit equations” \( z = f(x, y) \). That gives one \( z \) for each \( x \) and \( y \). A more general equation is \( F(x, y, z) = 0 \). An example is the sphere \( F = x^2 + y^2 + z^2 - 4 = 0 \). We could solve to find \( z = \sqrt{4 - x^2 - y^2} \) and also \( z = -\sqrt{4 - x^2 - y^2} \). These are two surfaces of the type \( z = f(x, y) \), to give the top half and bottom half of the sphere. In other examples it is difficult or impossible to solve for \( z \) and we really want to stay with the “implicit equation” \( F(x, y, z) = 0 \).

How do you find tangent planes and linear approximations for \( F(x, y, z) = 0 \)? Problem 3 shows by example.

3. The surface \( zz + 2yz - 10 = 0 \) goes through the point \((x_0, y_0, z_0) = (1, 2, 2)\). Find the tangent plane and normal vector. Estimate \( z \) when \( x = 1.1 \) and \( y = 1.9 \).
Main idea: Go ahead and differentiate \( F = xz + 2yz - 10 \). Not only \( x \) and \( y \) derivatives, also \( z \):

\[
\frac{\partial F}{\partial x} = z = 2 \quad \text{and} \quad \frac{\partial F}{\partial y} = 2z = 4 \quad \text{and} \quad \frac{\partial F}{\partial z} = x + 2y = 5 \quad \text{at the basepoint} \ (1, 2, 2).
\]

The tangent plane is \( 2(x - 1) + 4(y - 2) + 5(z - 2) = 0 \). The normal vector is \( N = (2, 4, 5) \). Notice how \( F_x, F_y, \) and \( F_z \) multiply \( \Delta x \) and \( \Delta y \) and \( \Delta z \). The total change is \( \Delta F \) which is zero (because \( F \) is constant: the surface is \( F = 0 \)). A linear approximation stays on the tangent plane! So if you know \( x = 1.1 \) and \( y = 1.9 \) you can solve for \( z \) on the plane:

\[
2(1.1 - 1) + 4(1.9 - 2) + 5(z - 2) = 0 \quad \text{gives} \quad z = \frac{2(1.1) - 4(1.9)}{5} = -\frac{2}{5}.
\]

I would memorize the tangent plane formula, which is

\[
F_x(x - x_0) + F_y(y - y_0) + F_z(z - z_0) = 0.
\]

In this example you could solve \( F = xz + 2yz - 10 = 0 \) to find \( z \). The explicit equation \( z = f(x, y) \) is

\[
F_x, F_y, \text{and } F_z, \text{ multiply } Ax \text{ and } Ay \text{ and } Az. \text{ The total change is } \Delta F \text{ which is zero (because } F \text{ is constant: the surface is } F = 0). \text{ A linear approximation stays on the tangent plane! So if you know } x = 1.1 \text{ and } y = 1.9 \text{ you can solve for } z \text{ on the plane:}
\]

\[
2(1.1 - 1) + 4(1.9 - 2) + 5(z - 2) = 0 \quad \text{gives} \quad z = \frac{2(1.1) - 4(1.9)}{5} = -\frac{2}{5}.
\]

The last topic in this important section is Newton's method. It deals with two functions \( g(x, y) \) and \( h(x, y) \). Solving \( g(x, y) = 0 \) should give a curve, solving \( h(x, y) = 0 \) should give another curve, and solving both equations should give the point (or points) where the two curves meet. When the functions are complicated—they usually are—we "linearize." Instead of \( g(x, y) = 0 \) and \( h(x, y) = 0 \) Newton solves

\[
\frac{\partial g}{\partial x} + \frac{\partial h}{\partial y} = 0
\]

Those are linear equations for \( Ax \) and \( Ay \). We move to the new basepoint \( (x_1, y_1) = (x_0 + \Delta x, y_0 + \Delta y) \) and start again. Newton's method solves many linear equations instead of \( g(x, y) = 0 \) and \( h(x, y) = 0 \).

4. Take one Newton step from \( (x_0, y_0) = (1, 2) \) toward the solution of \( g = xy - 3 = 0 \) and \( h = x + y - 2 = 0 \).

- The partial derivatives at the basepoint \( (1, 2) \) are \( g_x = y = 2 \) and \( g_y = x = 1 \) and \( h_x = 1 \) and \( h_y = 1 \). The functions themselves are \( g = -1 \) and \( h = 1 \). Newton solves the two linear equations above (tangent equations) for \( \Delta x \) and \( \Delta y \):

\[
\begin{align*}
-1 + 2\Delta x + \Delta y &= 0 \\
1 + \Delta x + \Delta y &= 0
\end{align*}
\]

\[
\Delta x = 2 \quad \Delta y = -3 \quad \text{The new guess is} \quad x_1 = x_0 + \Delta x = 3 \quad y_1 = y_0 + \Delta y = -1.
\]

The new point \( (3, -1) \) exactly solves \( h = x + y - 2 = 0 \). It misses badly on \( g = xy - 3 = 0 \). This surprised me because the method is usually terrific. Then I tried to solve the equations exactly by algebra.

Substituting \( y = 2 - x \) from the second equation into the first gave \( x(2 - x) - 3 = 0 \). This is a quadratic \( x^2 - 2x + 3 = 0 \). But it has no real solutions! Both roots are complex numbers. Newton never had a chance.

**Read-throughs and selected even-numbered solutions:**

The tangent line to \( y = f(x) = y - y_0 = f'(x_0)(x - x_0) \). The tangent plane to \( w = f(x, y) = w - w_0 = (\partial f/\partial x)_0(x - x_0) + (\partial f/\partial y)_0(y - y_0) \). The normal vector is \( N = (f_x, f_y, -1) \). For \( w = x^3 + y^3 \) the tangent equation at \( (1, 1, 2) \) is \( w - 2 = 3(x - 1) + 3(y - 1) \). The normal vector is \( N = (3, 3, -1) \). For a sphere, the direction of \( N \) is out from the origin.

The surface given implicitly by \( F(x, y, z) = c \) has tangent plane with equation \( (\partial F/\partial x)_0(x - x_0) + (\partial F/\partial y)(y - y_0) + (\partial F/\partial z)_0(z - z_0) = 0 \). For \( xz^2 = 6 \) at \( (1, 2, 3) \) the tangent plane has the equation
6(x - 1) + 3(y - 2) + 2(z - 3) = 0. On that plane the differentials satisfy 6dx + 3dy + 2dz = 0. The differential of z = f(x, y) is dz = f_x dx + f_y dy. This holds exactly on the tangent plane, while Δz ≈ f_x Δx + f_y Δy holds approximately on the surface. The height z = 3x + 7y is more sensitive to a change in y than in x, because the partial derivative ∂z/∂y = 7 is larger than ∂z/∂x = 3.

The linear approximation to f(x, y) is f(x_0, y_0) + (af/ax)(x - x_0) + (af/ay)(y - y_0). This is the same as Δf ≈ (af/ax)Δx + (af/ay)Δy. The error is of order (Δx)^2 + (Δy)^2. For f = sin xy the linear approximation around (0,0) is f_L = 0. We are moving along the tangent plane instead of the surface. When the equation is given as F(x, y, z) = c, the linear approximation is F_x Δx + F_y Δy + F_z Δz = 0.

Newton's method solves g(x, y) = 0 and h(x, y) = 0 by a linear approximation. Starting from x_n, y_n, the equations are replaced by g_x Δx + g_y Δy = -g(x_n, y_n) and h_x Δx + h_y Δy = -h(x_n, y_n). The steps Δx and Δy go to the next point (x_n + 1, y_n + 1). Each solution has a basin of attraction. Those basins are likely to be fractals.

8 \mathbf{N} = 8\pi i + 4\pi j - k; 8\pi(r - 2) + 4\pi(h - 2) = V - 8\pi

12 \mathbf{N}_1 = 2i + 4j - k \text{ and } \mathbf{N}_2 = 2i + 6j - k \text{ give } \mathbf{v} = \begin{vmatrix} i & j & k; 2 & 4 & -1 \end{vmatrix} = 2i + 4k \text{ tangent to both surfaces}

14 The direction of \mathbf{N} is 2xy^2i + 2x^2yj - k = 8i + 4j - k. So the line through (1,2,4) has x = 1 + 8t, y = 2 + 4t, z = 4 - t.

18 df = yz dx + xz dy + xy dz.

32 \frac{3}{4} Δx - Δy = \frac{3}{8} \text{ and } -Δx + \frac{3}{4} Δy = \frac{3}{8} \text{ give } Δx = Δy = -\frac{3}{2}. The new point is (-1, -1), an exact solution.

The point (\frac{1}{2}, \frac{1}{2}) is in the gray band (upper right in Figure 13.11a) or the blue band on the front cover.

38 A famous fractal shows the three basins of attraction – see almost any book on fractals. Remarkable property of the boundaries points between basins: they touch all three basins! Try to draw 3 regions with this property.

13.4 Directional Derivatives and Gradients

The partial derivatives ∂f/∂x and ∂f/∂y are directional derivatives, in special directions. They give the slope in directions u = (1, 0) and u = (0, 1), parallel to the x and y axes. From those two partial derivatives we can quickly find the derivative in any other direction u = (cos θ, sin θ):

directional derivative \quad D_uf = (\frac{∂f}{∂x})\cos θ + (\frac{∂f}{∂y})\sin θ.

It makes sense that the slope of the surface z = f(x,y), climbing at an angle between the x direction and y direction, should be a combination of slopes ∂f/∂x and ∂f/∂y. That slope formula is really a dot product between the direction vector u and the derivative vector (called the gradient):

gradient = \left(\frac{∂f}{∂x}, \frac{∂f}{∂y}\right) = \nabla f \quad \text{direction} = (\cos θ, \sin θ) = u \quad \text{directional derivative} = \nabla f \cdot u.
1. Find the gradient of \( f(x, y) = 4x + y - 7 \). Find the derivative in the 45° direction, along the line \( y = x \).
   - The partial derivatives are \( f_x = 4 \) and \( f_y = 1 \). So the gradient is the vector \( \nabla f = (4, 1) \).
   - Along the 45° line \( y = x \), the direction vector is \( \mathbf{u} = (\cos \frac{\pi}{4}, \sin \frac{\pi}{4}) \). This is \( \mathbf{u} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \). The dot product \( \nabla f \cdot \mathbf{u} = 4 \times \frac{\sqrt{2}}{2} + 1 \times \frac{\sqrt{2}}{2} = 5 \) is \( D_\mathbf{u} f \), the directional derivative.

2. Which direction gives the largest value of \( D_\mathbf{u} f \)? This is the steepest direction.
   - The derivative is the dot product of \( \nabla f = (4, 1) \) with \( \mathbf{u} = (\cos \theta, \sin \theta) \). A dot product equals the length \( |\nabla f| = \sqrt{4^2 + 1^2} = \sqrt{17} \) times the length \( |\mathbf{u}| = 1 \) times the cosine of the angle between \( \nabla f \) and \( \mathbf{u} \). To maximize the dot product and maximize that cosine, choose \( \mathbf{u} \) in the same direction as \( \nabla f \). Make \( \mathbf{u} \) a unit vector:
     \[
     \mathbf{u} = \frac{\nabla f}{|\nabla f|} = \left( \frac{4}{\sqrt{17}}, \frac{1}{\sqrt{17}} \right) \quad \text{and} \quad \nabla f \cdot \mathbf{u} = 4 \left( \frac{4}{\sqrt{17}} \right) + 1 \left( \frac{1}{\sqrt{17}} \right) = \frac{17}{\sqrt{17}} = \sqrt{17}.
     \]
   This is the general rule: The steepest direction is parallel to the gradient \( \nabla f = (f_x, f_y) \). The steepness (the slope) is \( |\nabla f| = \sqrt{f_x^2 + f_y^2} \). This is the largest value of \( D_\mathbf{u} f \).

3. Find a function \( f(x, y) \) for which the steepest direction is the x direction.
   - The question is asking for \( \frac{\partial f}{\partial y} = 0 \). Then the gradient is \( (\frac{\partial f}{\partial x}, 0) \). It points in the x-direction. The maximum slope is \( \sqrt{\left( \frac{\partial f}{\partial x} \right)^2 + 0^2} \) which is just \( |\frac{\partial f}{\partial x}| \).
   The answer is: Don't let \( f \) depend on \( y \). Choose \( f = x \) or \( f = e^x \) or any \( f(x) \). The slope in the y-direction is zero! The steepest slope is in the pure x-direction. At every in-between direction the slope is a mixture of \( \frac{\partial f}{\partial x} \) and 0. The steepest slope is \( |\frac{\partial f}{\partial x}| \) with no zero in the mixture.

\( \nabla f \cdot \mathbf{u} \) is the directional derivative along a straight line (in the direction \( \mathbf{u} \)). What if we travel along a curve? The value of \( f(x, y) \) changes as we travel, and calculus asks how fast it changes. This is an "instantaneous" question, at a single point on the curved path. At each point the path direction is the tangent direction. So replace the fixed vector \( \mathbf{u} \) by the tangent vector \( \mathbf{T} \) at that point: Slope of \( f(x, y) \) going along path = \( \nabla f \cdot \mathbf{T} \).

The tangent vector \( \mathbf{T} \) was in Section 12.1. We are given \( x(t) \) and \( y(t) \), the position as we move along the path. The derivative \( (\frac{dx}{dt}, \frac{dy}{dt}) \) is the velocity vector \( \mathbf{v} \). This is along the tangent direction (parallel to \( \mathbf{T} \)), but \( \mathbf{T} \) is required to be a unit vector. So divide \( \mathbf{v} \) by its length which is the speed \( |\mathbf{v}| = \sqrt{\left( \frac{dx}{dt} \right)^2 + \left( \frac{dy}{dt} \right)^2} = ds/dt \):
     \[
     \mathbf{v} = \left( \frac{dx}{dt}, \frac{dy}{dt} \right) \quad \text{gives} \quad \nabla f \cdot \mathbf{v} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}. \quad \text{This is} \quad \frac{df}{dt}.
     \]
     \[
     \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{v}}{ds/dt} = \mathbf{T} \quad \text{gives} \quad \nabla f \cdot \mathbf{T}. \quad \text{This is} \quad \frac{df}{ds}. \quad \text{This is} \quad \frac{df}{ds} = \frac{df}{dt} \frac{dt}{ds}.
     \]
   The speed is divided out of the slope \( df/ds \). The speed is not divided out of the rate of change \( df/dt \). One says how steeply you climb. The other says how fast you climb.

4. How steeply do you climb and how fast do you climb on a roller-coaster of height \( f(x, y) = 2x + y \)? You travel around the circle \( x = \cos 4t, \ y = \sin 4t \) with velocity \( \mathbf{v} = (-4 \sin 4t, 4 \cos 4t) \) and speed \( |\mathbf{v}| = 4 \).
   - The gradient of \( f = 2x + y \) is \( \nabla f = (2, 1) \). The tangent vector is \( \mathbf{T} = \frac{\mathbf{v}}{|\mathbf{v}|} = (-\sin 4t, \cos 4t) \).
     \[
     \text{Slope of path} = \nabla f \cdot \mathbf{T} = -2 \sin 4t + \cos 4t \quad \text{Maximum slope} \sqrt{5}.
     \]
     \[
     \text{Climbing rate} = \nabla f \cdot \mathbf{v} = -8 \sin 4t + 4 \cos 4t \quad \text{Maximum rate} \ 4\sqrt{5}.
     \]
   How fast you climb = (how steeply you climb) \times (how fast you travel).
Read-throughs and selected even-numbered solutions:

\( D_u f \) gives the rate of change of \( f(x, y) \) in the direction \( u \). It can be computed from the two derivatives \( \frac{\partial f}{\partial x} \) and \( \frac{\partial f}{\partial y} \) in the special directions \((1,0)\) and \((0,1)\). In terms of \( u_1, u_2 \) the formula is \( D_u f = f_x u_1 + f_y u_2 \). This is a dot product of \( u \) with the vector \((f_x, f_y)\), which is called the gradient. For the linear function \( f = ax + by \), the gradient is \( \text{grad } f = (a, b) \) and the directional derivative is \( D_u f = (a, b) \cdot u \).

The gradient \( \nabla f = (f_x, f_y) \) is not a vector in three dimensions, it is a vector in the base plane. It is perpendicular to the level lines. It points in the direction of steepest climb. Its magnitude \( |\text{grad} f| \) is the steepness \( \sqrt{f_x^2 + f_y^2} \). For \( f = x^2 + y^2 \) the gradient points out from the origin and the slope in that steepest direction is \( |(2x, 2y)| = 2r \).

The gradient of \( f(x, y, z) \) is \((f_x, f_y, f_z)\). This is different from the gradient on the surface \( F(x, y, z) = 0 \), which is \(-\frac{F_x}{F_z} \mathbf{i} - \frac{F_y}{F_z} \mathbf{j} \). Traveling with velocity \( v \) on a curved path, the rate of change of \( f \) is \( \frac{df}{dt} = (\text{grad } f) \cdot v \). When the tangent direction is \( T \), the slope of \( f \) is \( \frac{df}{ds} = (\text{grad } f) \cdot T \). In a straight direction \( u \), \( \frac{df}{ds} \) is the same as the directional derivative \( D_u f \).

12 In one dimension the gradient of \( f(x) \) is \( \frac{df}{dx} \mathbf{i} \). The two possible directions are \( u = \mathbf{i} \) and \( u = -\mathbf{i} \). The two directional derivatives are \( +\frac{df}{dx} \) and \( -\frac{df}{dx} \). The normal vector \( \mathbf{N} \) is \( \frac{df}{dx} \mathbf{i} - \mathbf{j} \).

14 Here \( f = 2x \) above the line \( y = 2x \) and \( f = y \) below that line. The two pieces agree on the line. Then \( \text{grad } f = 2\mathbf{i} \) above and \( \text{grad } f = \mathbf{j} \) below. Surprisingly \( f \) increases fastest along the line, which is the direction \( \mathbf{u} = \frac{1}{\sqrt{3}} (\mathbf{i} + 2\mathbf{j}) \) and gives \( D_u f = -\frac{2}{\sqrt{3}} \).

28 (a) False because \( f + C \) has the same gradient as \( f \) (b) True because the line direction \((1,1,-1)\) is also the normal direction \( \mathbf{N} \) (c) False because the gradient is in 2 dimensions.

30 \( \theta = \tan^{-1} \frac{y}{x} \) has \( \text{grad } \theta = \left( \frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) \). The unit vector in this direction is \( \mathbf{T} = \left( -\frac{y}{\sqrt{x^2 + y^2}}, \frac{x}{\sqrt{x^2 + y^2}} \right) \). Then \( \text{grad } \theta \cdot \mathbf{T} = \frac{y^2 + x^2}{x^2 + y^2} = \frac{1}{r} \).

34 The gradient is \((2ax + c)i + (2by + d)j\). The figure shows \( c = 0 \) and \( d \approx \frac{1}{3} \) at the origin. Then \( b \approx \frac{1}{3} \) from the gradient at \((0,1)\). Then \( a \approx -\frac{1}{4} \) from the gradient at \((2,0)\). The function \(-\frac{1}{4}x^2 + \frac{1}{3}y^2 + \frac{1}{3}y \) has hyperbolas opening upwards as level curves.

44 \( v = (2t, 0) \) and \( \mathbf{T} = (1,0) \); \( \text{grad } f = (y, x) \) so \( \frac{df}{dt} = 2ty = 6t \) and \( \frac{df}{ds} = y = 3 \).

48 \( D^2 = (x-1)^2 + (y-2)^2 \) has \( 2D \frac{\partial D}{\partial z} = 2(x-1) \) or \( \frac{\partial D}{\partial z} = \frac{x-1}{D} \). Similarly \( 2D \frac{\partial D}{\partial y} = 2(y-2) \) and \( \frac{\partial D}{\partial y} = \frac{y-2}{D} \).

Then \( |\text{grad } D| = \left( \frac{x-1}{D} \right)^2 + \left( \frac{y-2}{D} \right)^2 \) = 1. The graph of \( D \) is a 45° cone with its vertex at \((1,2)\).
13.5 The Chain Rule

Chain Rule 1  On the surface \( z = g(x, y) \) the partial derivatives of \( f(z) \) are \( \frac{\partial f}{\partial x} = \frac{df}{dx} \frac{\partial x}{\partial z} \) and \( \frac{\partial f}{\partial y} = \frac{df}{dy} \frac{\partial y}{\partial z} \). \( z = x^2 + y^2 \) gives a bowl. Then \( f(z) = \sqrt{z} = \sqrt{x^2 + y^2} \) gives a sharp-pointed cone. The slope of the cone in the \( z \)-direction is

\[
\frac{\partial f}{\partial z} = \frac{df}{dz} \frac{\partial z}{\partial x} = \frac{1}{2}z^{-1/2}(2x) = \frac{x}{\sqrt{z}} = \frac{x}{\sqrt{x^2 + y^2}}.
\]

Check that by directly taking the \( x \)-derivative of \( f(g(x, y)) \).

Chain Rule 2  For \( z = f(x, y) \) on the curve \( x = x(t) \) and \( y = y(t) \) the \( t \)-derivative is \( \frac{\partial z}{\partial t} = \frac{df}{dx} \frac{dx}{dt} + \frac{df}{dy} \frac{dy}{dt} \).

This is exactly the climbing rate from the previous section 13.4.

Chain Rule 3  For \( z = f(x, y) \) when \( x = x(t, u) \) and \( y = y(t, u) \) the \( t \)-derivative is \( \frac{\partial z}{\partial t} = \frac{df}{dx} \frac{dx}{dt} + \frac{df}{dy} \frac{dy}{dt} \).

This combines Rule 1 and Rule 2. The outer function \( f \) has two variables \( x, y \) as in Rule 2. The inner functions \( x \) and \( y \) have two variables as in Rule 1. So all derivatives are partial derivatives. But notice:

\[
\frac{\partial z}{\partial x} \frac{\partial x}{\partial u} \neq \frac{\partial z}{\partial u} \frac{\partial x}{\partial x}.
\]

The correct rule is

\[
\frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}.
\]

1. A change in \( u \) produces a change in \( x = tu \) and \( y = t/u \). These produce a change in \( z = 3x + 2y \). Find \( \frac{\partial z}{\partial u} \).

\[
\frac{\partial z}{\partial u} = \frac{3}{u} + 2\left(\frac{-t}{u^2}\right).
\]

2. When would Rule 3 reduce to Rule 2?  The inner functions \( x \) and \( y \) depend only on \( t \), not \( u \).

Please read the paradox on page 501. Its main point is: For partial derivatives you must know which variable is moving and also which variable is not moving.

Read-throughs and selected even-numbered solutions:

The chain rule applies to a function of a function. The \( x \)-derivative of \( f(g(x, y)) \) is \( \frac{df}{dx} \frac{dg}{dx} \frac{dx}{dz} \frac{dz}{dx} \) where \( 
\frac{\partial f}{\partial x} \frac{dx}{dz} + \frac{\partial f}{\partial y} \frac{dy}{dz} \). The example \( f = (x+y)^n \) has \( g = x + y \). Because \( \partial g/\partial x = \partial g/\partial y \) we know that \( \partial f/\partial x = \partial f/\partial y \). This partial differential equation is satisfied by any function of \( x + y \).

Along a path, the derivative of \( f(x(t), y(t)) \) is \( df/dt = (\partial f/\partial x)(dx/dt) + (\partial f/\partial y)(dy/dt) \). The derivative of \( f(x(t), y(t), z(t)) \) is \( f_x x_t + f_y y_t + f_z z_t \). If \( f = xy \) then the chain rule gives \( df/dt = y \, dx/dt + x \, dy/dt \). That is the same as the product rule! When \( x = u \, t \) and \( y = u^2 t \) the path is a straight line. The chain rule for \( f(x, y) \) gives \( df/dt = f_x u_1 + f_y u_2 \). That is the directional derivative \( D_u f \).

The chain rule for \( f(x(t, u), y(t, u)) \) is \( \partial f/\partial t = (\partial f/\partial x)(\partial x/\partial t) + (\partial f/\partial y)(\partial y/\partial t) \). We don't write \( df/\partial t \) because \( f \) also depends on \( u \). If \( z = r \cos \theta \) and \( y = r \sin \theta \), the variables \( t, u \) change to \( r \) and \( \theta \). In this case \( \partial f/\partial r = (\partial f/\partial x) \cos \theta + (\partial f/\partial y) \sin \theta \) and \( \partial f/\partial \theta = (\partial f/\partial x)(-r \sin \theta) + (\partial f/\partial y)(r \cos \theta) \). That connects the derivatives in rectangular and polar coordinates. The difference between \( \partial r/\partial x = x/r \) and \( \partial r/\partial z = 1/cos \theta \) is because \( y \) is constant in the first and \( \theta \) is constant in the second.

With a relation like \( xyz = 1 \), the three variables are not independent. The derivatives \( (\partial f/\partial x)_y \) and \( (\partial f/\partial x)_z \) and \( (\partial f/\partial x) \) mean that \( y \) is held constant, and \( z \) is constant, and both are constant. For
13.6 Maxima, Minima, and Saddle Points  

A one-variable function \( f(x) \) reaches its maximum and minimum at three types of critical points:

1. Stationary points where \( \frac{df}{dx} = 0 \)
2. Rough points
3. Endpoints (possibly at \( \infty \) or \( -\infty \)).

A two-variable function \( f(x, y) \) has the same three possible types of critical points:

1. Stationary points where \( \frac{\partial f}{\partial x} = 0 \) and \( \frac{\partial f}{\partial y} = 0 \)
2. Rough points
3. Boundary points.

The stationary points come first. Notice that they involve two equations (both partial derivatives are zero).

There are two unknowns (the coordinates \( x \) and \( y \) of the stationary point). The tangent is horizontal as usual, but it is a tangent plane to the surface \( z = f(x, y) \).

It is harder to solve two equations than one. And the second derivative test (which was previously \( f'' > 0 \) for a minimum and \( f'' < 0 \) for a maximum) now involves all three derivatives \( f_{xx}, f_{yy}, \) and \( f_{xy} = f_{yx} \):

- **Minimum**  
  \[ f_{xx} > 0 \quad f_{xx}f_{yy} > (f_{xy})^2 \]
  - **Maximum**  
  \[ f_{xx} < 0 \quad f_{xx}f_{yy} > (f_{xy})^2 \]
  - **Saddle**  
  \[ f_{xx}f_{yy} < (f_{xy})^2 \]

When \( f_{xx}f_{yy} = (f_{xy})^2 \) the test gives no answer. This is like \( f'' = 0 \) for a one-variable function \( f(x) \).

Our two-variable case really has a 2 by 2 matrix of second derivatives. Its determinant is the critical quantity \( f_{xx}f_{yy} - (f_{xy})^2 \). This pattern continues on to \( f(x, y, z) \) or \( f(x, y, z, t) \). Those have 3 by 3 and 4 by 4 matrices of second derivatives and we check 3 or 4 determinants. In linear algebra, a positive definite second-derivative matrix indicates that the stationary point is a minimum.

1. \((13.6.26)\) Find the stationary points of \( f(x, y) = xy - \frac{1}{4}x^4 - \frac{1}{4}y^4 \) and decide between min, max, and saddle.
The partial derivatives are $f_x = y - x^3$ and $f_y = x - y^3$. Set both derivatives to zero: 

$$y = x^3$$ and $$x = y^3.$$ This gives $y = 0, 1,$ or $-1$. Then $x = y^3$ gives $x = 0, 1,$ or $-1$.

The stationary points are $(0,0)$ and $(1,1)$ and $(-1,-1)$. The second derivatives are $f_{xx} = -3x^2$ and $f_{yy} = -3y^2$ and $f_{xy} = 1$:

- $(0,0)$ is a saddle point because $f_{xx}f_{yy} = (0)(0)$ is less than $(1)^2$.
- $(1,1)$ and $(-1,-1)$ are maxima because $f_{xx}f_{yy} = (-3)(-3)$ is greater than $(1)^2$ and $f_{xx} = -3$.

Solving $x = y^3$ and $y = x^3$ is our example of the two-variable Newton method in Section 11.3. This is really important in practice. For this function $xy - \frac{1}{4}x^4 - \frac{1}{4}y^4$ we found a saddle point and two maximum points. The minimum is at infinity. This counts as a "boundary point".

2. (This is Problem 13.6.56) Show that a solution to Laplace’s equation $f_{xx} + f_{yy} = 0$ has no maximum or minimum stationary points. So where are the maximum and minimum of $f(x, y)$?

- A maximum requires $f_{xx} < 0$. It also requires $f_{yy} < 0$. We didn’t say that, but it follows from the requirement $f_{xx}f_{yy} > (f_{xy})^2$. The left side has to be positive, so $f_{xx}$ and $f_{yy}$ must have the same sign.

If $f_{xx} + f_{yy} = 0$ this can’t happen; stationary points must be saddle points (or $f = constant$). A max or min is impossible. Those must occur at rough points or boundary points.

Example A $f(x,y) = \ln(x^2 + y^2)$ has a minimum of $-\infty$ at $(x,y) = (0,0)$, since $\ln 0 = -\infty$. This is a rough point because $f_x = \frac{2x}{x^2+y^2}$ is unbounded. You could check Laplace’s equation two ways. One is to compute $f_{xx} = \frac{2}{x^2+y^2} - \left(\frac{2x}{x^2+y^2}\right)^2$. Also $f_{yy} = \frac{2}{x^2+y^2} - \left(\frac{2y}{x^2+y^2}\right)^2$. Add to get zero. The other way is to write $f = \ln r^2 = 2\ln r$ in polar coordinates. Then $f_r = \frac{2}{r}$ and $f_{rr} = -\frac{2}{r^2}$. Substitute into the polar Laplace equation to get $f_{rr} + \frac{1}{r}f_r + f_{\theta\theta} = 0$.

Example B $f(x,y) = xy$ satisfies Laplace’s equation because $f_{xx} + f_{yy} = 0 + 0$. The stationary point at the origin cannot be a max or min. It is a typical and famous saddle point: We find $f_{xy} = 1$ and then $f_{xx}f_{yy} = (0)(0) < (1)^2$. There are no rough points. The min and max must be at boundary points.

Note: Possibly there are no restrictions on $x$ and $y$. The boundary is at infinity. Then the max and min occur out at infinity. Maximum when $x$ and $y$ go to $+\infty$. Minimum when $x \to +\infty$ and $y \to -\infty$, because then $xy \to -\infty$. (Also max when $x$ and $y$ go to $-\infty$. Also min when $x \to -\infty$ and $y \to +\infty$).

Suppose $x$ and $y$ are restricted to stay in the square $1 \leq x \leq 2$ and $1 \leq y \leq 2$. Then the max and min of $xy$ occur on the boundary of the square. Maximum at $x = y = 2$. Minimum at $x = y = 1$. In a way those are “rough points of the boundary,” because they are sharp corners.

Suppose $x$ and $y$ are restricted to stay in the unit circle $x = \cos t$ and $y = \sin t$. The maximum of $xy$ is on the boundary (where $xy = \cos t \sin t$). The circle has no rough points. The maximum is at the $45^\circ$ angle $t = \frac{\pi}{4}$ (also at $t = \frac{5\pi}{4}$). At those points $xy = \cos \frac{\pi}{4} \sin \frac{\pi}{4} = \frac{1}{2}$. To emphasize again: This maximum occurred on the boundary of the circle.
Finally we call attention to the Taylor Series for a function $f(x, y)$. The text chose $(0,0)$ as basepoint. The whole idea is to match each derivative $\frac{\partial^k}{\partial x^k} f(x, y)$ at the basepoint by one term in the Taylor series. Since $\frac{\partial^k}{\partial x^k} f(x, y)$ has derivative equal to 1, multiply this standard power by the required derivative to find the correct term in the Taylor Series.

When the basepoint moves to $(x_0, y_0)$, change from $x^n y^m$ to $(x - x_0)^n (y - y_0)^m$. Divide by the same $n!m!$

3. Find the Taylor series of $f(x, y) = e^{x-y}$ with $(0,0)$ as the basepoint. Notice $f(x, y) = e^x$ times $e^{-y}$.

- Method 1: Multiply the series for $e^x$ and $e^{-y}$ to get $e^{x-y} = (1 + x + \frac{1}{2!} x^2 + \cdots)(1 - y + \frac{1}{2!} y^2 + \cdots) = 1 + x - y + \frac{1}{2} x^2 - xy + \cdots$

- Method 2: Substitute $x - y$ directly into the series to get $e^{x-y} = 1 + (x - y) + \frac{1}{2!}(x - y)^2 + \cdots.$

- Method 3: (general method): Find all the derivatives of $f(x, y) = e^{x-y}$ at the basepoint $(0,0)$:

$$f(0,0) = 1 \quad f_x(0,0) = 1 \quad f_y(0,0) = -1 \quad f_{xx}(0,0) = 1 \quad f_{xy}(0,0) = -1 \quad f_{yy}(0,0) = 1 \quad \cdots$$

Then the Taylor Series is $\frac{1}{0!1!} + \frac{1}{1!1!}x + \frac{-1}{1!1!}y + \frac{1}{2!0!}x^2 + \frac{-1}{1!1!}xy + \frac{1}{1!1!}y^2 + \cdots$. Remember that $0! = 1$.

Read-throughs and selected even-numbered solutions:

A minimum occurs at a stationary point (where $f_x = f_y = 0$) or a rough point (no derivative) or a boundary point. Since $f = x^2 - xy + 2y$ has $f_x = 2x - y$ and $f_y = 2 - x$, the stationary point is $x = 2, y = 4$. This is not a minimum, because $f$ decreases when $y = 2x$ increases.

The minimum of $d^2 = (x - x_1)^2 + (y - y_1)^2$ occurs at the rough point $(x_1, y_1)$. The graph of $d$ is a cone and grad $d$ is a unit vector that points out from $(x_1, y_1)$. The graph of $f = |xy|$ touches bottom along the lines $x = 0$ and $y = 0$. Those are “rough lines” because the derivative does not exist. The maximum of $d$ and $f$ must occur on the boundary of the allowed region because it doesn’t occur inside.

When the boundary curve is $x = x(t), y = y(t)$, the derivative of $f(x, y)$ along the boundary is $f_x x_t + f_y y_t$ (chain rule). If $f = x^2 + 2y^2$ and the boundary is $x = \cos t, y = \sin t$, then $df/dt = 2 \sin t \cos t$. It is zero at the points $t = 0, \pi/2, \pi, 3\pi/2$. The maximum is at $(0, \pm 1)$ and the minimum is at $(\pm 1, 0)$. Inside the circle $f$ has an absolute minimum at $(0,0)$.

To separate maximum from minimum from saddle point, compute the second derivatives at a stationary point. The tests for a minimum are $f_{xx} > 0$ and $f_{xx} f_{yy} > f_{xy}^2$. The tests for a maximum are $f_{xx} < 0$ and $f_{xx} f_{yy} > f_{xy}^2$. In case $ac < b^2$ or $f_{xx} f_{yy} < f_{xy}^2$, we have a saddle point. At all points these tests decide between concave up and concave down and "indefinite". For $f = 8x^2 - 6xy + y^2$, the origin is a saddle point. The signs of $f$ at $(1,0)$ and $(1,3)$ are $+$ and $-$.

The Taylor series for $f(x, y)$ begins with the terms $f(0,0) + x f_x + y f_y + \frac{1}{2} x^2 f_{xx} + xy f_{xy} + \frac{1}{2} y^2 f_{yy}$. The coefficient of $x^n y^m$ is $\frac{\partial^n \partial^m f}{\partial x^n \partial y^m}(0,0)$ divided by $n!m!$ To find a stationary point numerically, use
Newton's method or steepest descent.

18 Volume is $xyz = xy(1 - 3x - 2y) = xy - 3x^2 - 2y^2$; $V_x = y - 6x - 2y^2$ and $V_y = x - 4y$; at $(0, \frac{1}{2}, 0)$ and $(\frac{1}{2}, 0, 0)$ and $(0, 0, 1)$ the volume is $V = 0$ (minimum); at $(\frac{1}{4}, \frac{12}{48}, \frac{21}{48})$ the volume is $V = \frac{7}{3072}$ (maximum).

22 $\frac{\partial f}{\partial x} = 2x + 2$ and $\frac{\partial f}{\partial y} = 2y + 4$. (a) Stationary point $(-1, -2)$ yields $f_{\text{min}} = -5$. (b) On the boundary $y = 0$ the minimum of $x^2 + 2$ is $1$ at $(-1, 0)$ and $(1, 0, 0)$ the minimum is $0$ at $(0, 0, 1)$ the volume is $V = 0$ (minimum); at $(x, y, z)$ the volume is $V = -m$ (maximum).

28 $d_1 = x, d_2 = d_3 = \sqrt{(1 - x)^2 + 1}, \frac{df}{dx}(x + 2\sqrt{(1 - x)^2 + 1}) = 1 + \frac{-2(x-1)}{\sqrt{(1-x)^2+1}} = 0$ when $(1 - x)^2 + 1 = 4(x - 1)^2$ or $1 - x = \frac{1}{\sqrt{3}}$ or $x = 1 - \frac{1}{\sqrt{3}}$. From that point to $(1, 1, 1)$ the line goes up $1$ and across a $60^\circ$ angle with the horizontal that confirms three $120^\circ$ angles.

32 From the point $C = (0, -4)$ the lines to $(-1, 0)$ and $(1, 0)$ make a $60^\circ$ angle. $C$ is the center of the circle $x^2 + (y - \frac{1}{2})^2 = \frac{1}{4}$ through those two points. From any point on that circle, the lines to $(-1, 0)$ and $(1, 0)$ make an angle of $2x = 60^\circ = 120^\circ$. Theorem from geometry: angle from circle $= 2x$ angle from center.

34 $f(x, y)$ is $x + y$ for $n = 0, e$ for $n = 1$, zero for $n > 1$. Taylor series $xe^y = x + xy + \frac{1}{2}xy^2 + \frac{1}{3!}xy^3 + \cdots$

40 $f(x, y, z) = z^2 + y^2 + z^2$ with the constraint $g(x, y, z) = ax + by + cz = d$. 

13.7 Constraints and Lagrange Multipliers

In reality, a constraint $g(x, y) = k$ is very common. The point $(x, y)$ is restricted to this curve, when we are minimizing or maximizing $f(x, y)$. (Not to the inside of the curve, but right on the curve.) It is like looking for a maximum at a boundary point. The great difficulty is that we lose the equations $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$.

The great success of Lagrange multipliers is to bring back the usual equations "$x$ derivative equals zero" and "$y$ derivative equals zero." But these are not $f_x$ and $f_y$. We must account for the constraint $g(x, y) = k$.

The idea that works is to subtract an unknown multiple $\lambda$ times $g(x, y) - k$. Now set derivatives to zero:

$$\frac{\partial}{\partial x}[f(x, y) - \lambda(g(x, y) - k)] = 0 \quad \text{or} \quad \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}$$

$$\frac{\partial}{\partial y}[f(x, y) - \lambda(g(x, y) - k)] = 0 \quad \text{or} \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}.$$ 

The text explains the reasoning that leads to these equations. Here we solve them for $x, y,$ and $\lambda$. That locates the constrained maximum or minimum.

1. (This is Problem 13.7.6) Maximize $f(x, y) = x + y$ subject to $g(x, y) = x^{1/3}y^{2/3} = 1$. That is a special case of the Cobb-Douglas constraint: $x^\alpha y^{1-\alpha} = 1$.

- $f_x = \lambda g_x$ is $1 = \lambda(\frac{1}{3}x^{-2/3}y^{2/3})$ and $f_y = \lambda g_y$ is $1 = \lambda(\frac{2}{3}x^{1/3}y^{-1/3})$. The constraint is $1 = x^{1/3}y^{2/3}$.

Square the second equation and multiply by the first to get $1 = \left(\frac{2}{3}\right)^2(\frac{1}{3})$ or $\left(\frac{1}{3}\right)^3 = \frac{1}{4}$ or $\frac{1}{3} = 4^{-1/3}$. Then divide the constraint by the first equation to get $1 = \frac{1}{9}x$ or $x = \frac{9}{4}$ or $x = 4^{1/3}$. Divide the constraint by the second equation to get $1 = \frac{9}{3}y$ or $y = \frac{3}{4} = 4^{-1/3}$. Divide the constraint by the second equation to get $1 = \frac{9}{3}y$ or $y = \frac{3}{4} = 4^{-1/3}$. The constrained maximum is $f = x + y = 3 \cdot 4^{-1/3}$.

2. (This is Problem 13.7.22 and also Problem 13.7.8 with a twist. It gives the shortest distance to a plane.) Minimize $f(x, y, z) = x^2 + y^2 + z^2$ with the constraint $g(x, y, z) = ax + by + cz = d$. 

190
Now we have three variables \( z, y, a \) (also \( X \) for the constraint). The method is the same:

\[
\begin{align*}
f_x &= \lambda f_x = 2x = \lambda a \\
f_y &= \lambda g_y = 2y = \lambda b \\
f_z &= \lambda g_z = 2z = \lambda c.
\end{align*}
\]

Put \( x = \frac{1}{2} \lambda a \) and \( y = \frac{1}{2} \lambda b \) and \( z = \frac{1}{2} \lambda c \) in the constraint to get \( \frac{1}{2} \lambda (a^2 + b^2 + c^2) = d \). That yields \( \lambda \).

The constrained minimum is \( x^2 + y^2 + z^2 = \left( \frac{1}{2} \lambda \right)^2 (a^2 + b^2 + c^2) = \frac{d^2}{a^2 + b^2 + c^2} \).

The shortest distance to the plane is the square root \( \sqrt{\frac{|d|}{\sqrt{a^2 + b^2 + c^2}}} \). This agrees with the formula \( \frac{|d|}{|N|} \) from Section 11.2, where the normal vector to the plane was \( N = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \).

The text explains how to handle two constraints \( g(x, y, z) = k_1 \) and \( h(z, y, a) = k_2 \). There are two Lagrange multipliers \( \lambda_1 \) and \( \lambda_2 \). The text also explains inequality constraints \( g(x, y) \leq k \). The point \( (z, y) \) is either on the boundary where \( g(x, y) = k \) or it is inside where \( g(x, y) < k \). We are back to our old problem:

The minimum of \( f(x, y) \) may be at a boundary point. Using Lagrange multipliers we find \( \lambda > 0 \).

The minimum of \( f(x, y) \) may be at a stationary point. Using Lagrange multipliers we find \( \lambda = 0 \).

The second case has an inside minimum. The equation \( f_x = \lambda g_x \) becomes \( f_x = 0 \). Similarly \( f_y = 0 \). Lagrange is giving us one unified way to handle stationary points (inside) and boundary points. Rough points are handled separately. Problems 15–18 develop part of the theory behind \( \lambda \). I am most proud of including what calculus authors seldom attempt – the meaning of \( \lambda \). It is the derivative of \( f_{\min} \) with respect to \( k \). Thus \( \lambda \) measures the sensitivity of the answer to a change in the constraint.

This section is not easy but it is really important. Remember it when you need it.

**Read-throughs and selected even-numbered solutions:**

A restriction \( g(x, y) = k \) is called a constraint. The minimizing equations for \( f(x, y) \) subject to \( g = k \) are \( \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \) and \( g = k \). The number \( \lambda \) is the Lagrange multiplier. Geometrically, grad \( f \) is parallel to grad \( g \) at the minimum. That is because the level curve \( f = f_{\min} \) is tangent to the constraint curve \( g = k \). The number \( \lambda \) turns out to be the derivative of \( f_{\min} \) with respect to \( k \). The Lagrange function is \( L = f(x, y) - \lambda (g(x, y) - k) \) and the three equations for \( x, y, \lambda \) are \( \frac{\partial L}{\partial x} = 0 \) and \( \frac{\partial L}{\partial y} = 0 \) and \( \frac{\partial L}{\partial \lambda} = 0 \).

To minimize \( f = x^2 - y \) subject to \( g = x - y = 0 \), the three equations for \( x, y, \lambda \) are \( 2x = \lambda, -1 = -\lambda, x - y = 0 \). The solution is \( x = \frac{1}{2}, y = \frac{1}{2}, \lambda = 1 \). In this example the curve \( f(x, y) = f_{\min} = -\frac{1}{4} \) is a parabola which is tangent to the line \( g = 0 \) at \( (x_{\min}, y_{\min}) \).

With two constraints \( g(x, y, z) = k_1 \) and \( h(z, y, a) = k_2 \) there are two multipliers \( \lambda_1 \) and \( \lambda_2 \). The five unknowns are \( x, y, z, \lambda_1, \) and \( \lambda_2 \). The five equations are \( f_x = \lambda_1 g_x + \lambda_2 h_x, f_y = \lambda_1 g_y + \lambda_2 h_y, f_z = \lambda_1 g_z + \lambda_2 h_z, g = 0, \) and \( h = 0 \). The level surface \( f = f_{\min} \) is tangent to the curve where \( g = k_1 \) and \( h = k_2 \). Then grad \( f \) is perpendicular to this curve, and so are grad \( g \) and grad \( h \). With nine variables and six constraints, there will be six multipliers and eventually 15 equations. If a constraint is an inequality \( g \leq k \), then its multiplier must satisfy \( \lambda \leq 0 \) at a minimum.

\[
2x^2 + y^2 = 1 \text{ and } 2xy = \lambda (2z) \text{ and } x^2 = \lambda (2y) \text{ yield } 2\lambda^2 + \lambda^2 = 1. \text{ Then } \lambda = \frac{1}{\sqrt{3}} \text{ gives } x_{\max} = \pm \frac{\sqrt{6}}{3}, \text{ and } y_{\max} = \frac{\sqrt{3}}{3}, f_{\max} = \frac{2\sqrt{3}}{9}. \text{ Also } \lambda = -\frac{1}{\sqrt{3}} \text{ gives } f_{\min} = -\frac{2\sqrt{3}}{9}.
\]
18 \( f = 2x + y = 1001 \) at the point \( x = 1000, y = -999 \). The Lagrange equations are \( 2 = \lambda \) and \( 1 = \lambda \) (no solution). Linear functions with linear constraints generally have no maximum.

20 (a) \( yz = \lambda, xz = \lambda, zx = \lambda, \) and \( x + y + z = k \) give \( z = y = z = \frac{k}{3} \) and \( \lambda = \frac{k^3}{9} \) (b) \( V_{\text{max}} = \left( \frac{k}{3} \right)^3 \) so \( \partial V_{\text{max}} / \partial k = k^2 / 9 \) (which is \( \lambda ! \)) (c) Approximate \( \Delta V = \lambda \) times \( \Delta k = \frac{108^2}{9} (111 - 108) = 3888 \) in\(^3\).

26 Reasoning: By increasing \( k \), more points satisfy the constraints. More points are available to minimize \( f \). Therefore \( f_{\text{min}} \) goes down.

28 \( \lambda = 0 \) when \( h > k \) (not \( h = k \)) at the minimum. Reasoning: An increase in \( k \) leaves the same minimum. Therefore \( f_{\text{min}} \) is unchanged. Therefore \( \lambda = df_{\text{min}} / dk \) is zero.

13 Chapter Review Problems

Graph Problems

G1 Draw the level curves of the function \( f(x, y) = y - x \). Describe the surface \( z = y - x \).

G2 Draw the level curves of \( f(x, y) = \frac{4}{x - \frac{3}{2}} \). Label the curve through (3,3). Which points \((x, y)\) are not on any level curve? The surface has an infinite crack like an asymptote.

Computing Problems

C1 Set up Newton's method to give two equations for \( \Delta x \) and \( \Delta y \) when the original equations are \( y = x^5 \) and \( z = y^5 \). Start from various points \((x_0, y_0)\) to see which solutions Newton converges to. Compare the basins of attraction to Figure 13.3 and the front cover of this Guide.

Review Problems

R1 For \( f(x, y) = x^n y^m \) find the partial derivatives \( f_x, f_y, f_{xx}, f_{xy}, f_{yx}, \) and \( f_{yy} \).

R2 If \( z(x, y) \) is defined implicitly by \( F(x, y, z) = yz - yz + z = 0 \), find \( \partial z / \partial x \) and \( \partial z / \partial y \).

R3 Suppose \( z \) is a function of \( x/y \). From \( z = f(x/y) \), show that \( x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0 \).

R4 Write down a formula for the linear approximation to \( z = f(x, y) \) around the origin. If \( f(x, y) = 9 + xy \) show that the linear approximation at \((1,1)\) gives \( f \approx 11 \) while the correct value is 10.

R5 Find the gradient vector for the function \( f(x, y) = xy^2 \). How is the direction of the gradient at the point \( x = 1, y = 2 \) related to the level curve \( xy^2 = 4 \)?

R6 Find the gradient vector in three dimensions for the function \( F(x, y, z) = z - x^2y^2 \). How is the direction of the gradient related to the surface \( z = x^2y^2 \)?

R7 Give a chain rule for \( df / dt \) when \( f = f(x, y, z) \) and \( x, y, z \) are all functions of \( t \).

R8 Find the maximum value of \( f(x, y) = x + 2y - x^2 + xy - y^2 \).

R9 The minimum of \( x^2 + y^2 \) occurs on the boundary of the region \( R \) (not inside) for which regions?

R10 To minimize \( x^2 + y^2 \) on the line \( x + 3y = k \), introduce a Lagrange multiplier \( \lambda \) and solve the three equations for \( x, y, \lambda \). Check that the derivative \( df_{\text{min}} / dk \) equals \( \lambda \).
Drill Problems

D1  If \( z = \ln \sqrt{x^2 + y^2} \) show that \( \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 1 \) and \( zz_x + z_{yy} = 0 \) except at ______.

D2  The equation of the tangent plane to \( z = x^2 + y^3 \) at (1,1,2) is ______.

D3  The normal vector to the surface \( xyz^2 = 1 \) at (1,1,1) is \( \mathbf{N} = ______ \).

D4  The linear approximation to \( x^2 + y^2 \) near the basepoint (1,2) is ______.

D5  Find the directional derivative of \( f(x, y) = xe^y \) at the point (2,2) in the 45° direction \( y = x \). What is \( u \)? Compare with the ordinary derivative of \( f(x) = xe^x \) at \( x = 2 \).

D6  What is the steepest slope on the plane \( z = x + 2y \)? Which direction is steepest?

D7  From the chain rule for \( f(x, y) = xy^2 \) with \( x = u + v \) and \( y = uv \) compute \( \frac{\partial L}{\partial u} \) at \( u = 2, v = 3 \). Check by taking the derivative of \((u + v)(uv)^2\).

D8  What equations do you solve to find stationary points of \( f(x, y) \)? What is the tangent plane at those points? How do you know from \( f_{xx}, f_{xy}, \) and \( f_{yy} \) whether you have a saddle point?

D9  Find two functions \( f(x, y) \) that have \( \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} \) at all points. Which is the steepest direction on the surface \( z = f(x, y) \)? Which is the level direction?

D10 If \( x = r \cos \theta \) and \( y = r \sin \theta \) compute the determinant \( J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \)

D11 If \( r = \sqrt{x^2 + y^2} \) and \( \theta = \tan^{-1} \frac{y}{x} \) compute the determinant \( J^* = \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} = \frac{1}{r} \).