Business Math II (part 1): Multivariable Optimization (by Evan Dummit, 2019, v. 1.01)

Contents

1 Multivariable Optimization **1**

1.1 Functions of Several Variables and 3-Space . . . 1

1.1.1 Graphing Functions of 2 Variables: Surfaces and Level Sets . . . 2

1.1.2 Graphing Functions of 3 Variables: Level Surfaces . . . 4

1.2 Partial Derivatives . . . 5

1.3 Local Extreme Points and Optimization . . . 7

1.3.1 Minima, Maxima, and Saddle Points . . . 7

1.3.2 Classifying Critical Points . . . 9

1.4 Optimization of a Function on a Region, Linear Programming . . . 11

1.4.1 Optimization on a Region . . . 11

1.4.2 Linear Programming . . . 13

1.5 Lagrange Multipliers and Constrained Optimization . . . 16

---

1 Multivariable Optimization

In this chapter, we introduce 3-dimensional space (often called 3-space for short) and functions of several variables. We then develop the theory of differentiation for functions of several variables, and discuss applications to optimization and finding local extreme points. The development of these topics is very similar to that of one-variable calculus, save for some additional complexities due to the presence of extra variables.

1.1 Functions of Several Variables and 3-Space

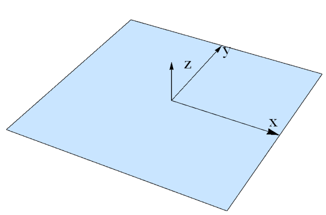
- A function of several variables is, as the name indicates, a function that takes in several variables and outputs a value associated to the inputs.
  - Example: The function $f(x, y) = x + y$ takes in two values x and y and outputs their sum $x + y$.
  - Example: The function $g(x, y, z) = x\,y\,z^2$ takes in three values x, y, and z, and outputs the product $x\,y\,z^2$.
  - Example: The function $d(l, w) = \sqrt{l^2 + w^2}$ gives the length of a diagonal of a rectangle as a function of the rectangle's length l and its width w.
  - Example: The function $V(r, h) = \frac{1}{3}\pi r^2 h$ gives the volume of a (right circular) cone as a function of its base radius r and its height h.
- As with functions of one variable, a function of several variables has a domain and a range: the domain is the set of input values and the range is the set of output values.
  - For a function of two variables $f(x, y)$, the domain is now a subset of the 2-dimensional plane rather than a subset of the real line. Because of this, domains of functions of more than one variable can be rather complicated.

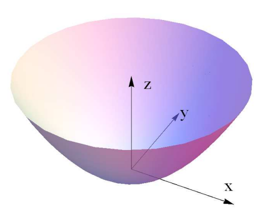
- In general, unless specified, the domain of a function is the largest possible set of inputs for which the definition of the function makes sense. We generally adopt the conventions that square roots of negative real numbers are not allowed, nor is division by zero.
- Example: For $f(x, y) = \sqrt{x} + \sqrt{y}$, the domain is the first quadrant of the xy-plane, defined by the inequalities $x \geq 0$ and $y \geq 0$.
- Example: For $f(x, y) = \sqrt{x^2 + y^2 - 1}$, the domain is the set of points in the xy-plane which satisfy $x^2 + y^2 \geq 1$: this describes all the points of the plane except for those lying strictly inside the unit circle.
- Example: For $f(x, y) = \dfrac{1}{x - y}$, the domain is the set of points in the xy-plane which satisfy $x - y \neq 0$: this describes all points in the plane except those on the line $y = x$.

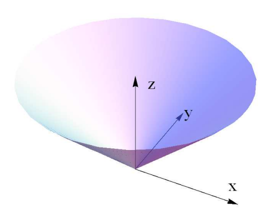
- Points in 3-space are represented by a triplet of numbers (x, y, z). The new coordinate z represents height above the xy-plane.
  - From an algebraic standpoint, 3-space behaves quite similarly to 2-dimensional space.
  - For example, in 3-space, we can measure the distance between any two points. By a suitable pair of applications of the Pythagorean Theorem, we can compute that the distance between points (x_1, y_1, z_1) and (x_2, y_2, z_2) is given by $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$.

1.1.1 Graphing Functions of 2 Variables: Surfaces and Level Sets

- There are two primary ways to visualize a function $f(x, y)$ of two variables.
- The first way is to plot the points (x, y, z) in 3-dimensional space satisfying $z = f(x, y)$.
  - At the point (x, y) in the plane, the graph has the height $z = f(x, y)$; so we see that as (x, y) varies through the plane, the function $z = f(x, y)$ will trace out a surface, called the graph of $f(x, y)$.
- Some example of simple graphs $z = f(x, y)$ are given below.
  - Example: The graph $z = 0$ is the xy-plane.
    - Note more generally that any equation of the form $a\,x + b\,y + c\,z = d$ for some constants a, b, c, d (with not all of a, b, c zero) will give a plane. We will study planes in more depth in a later section.
  - Example: The graph $z = x^2 + y^2$ is a paraboloid (i.e., a parabolic dish).
  - Example: The graph $z = \sqrt{x^2 + y^2}$ is a right circular cone opening upward, with vertex at the origin.



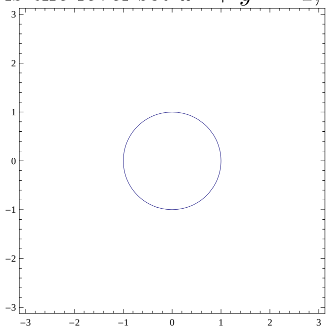
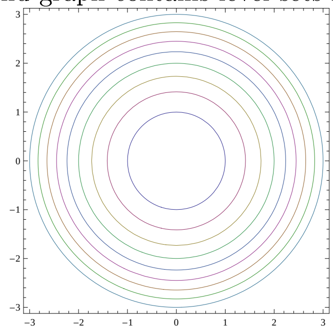




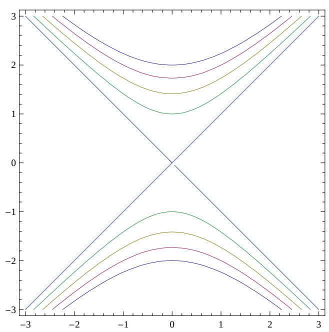
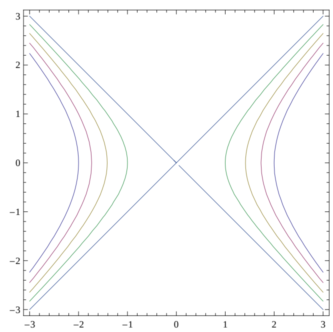
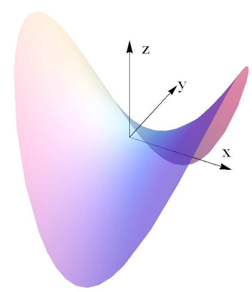
- The second way is to plot the points (x, y) in the plane on the level sets $f(x, y) = c$ (for particular values of c), as implicit curves.
  - For a given function $f(x, y)$ and a particular value of c, the points (x, y) satisfying $f(x, y) = c$ are called a level set of f. For a function of two variables these sets will generally be curves, so they are also sometimes called level curves.
  - If we graph many of these level curves together on the same axes, we will obtain a "topographical map" of the function $f(x, y)$.

- Some examples of level sets are given below:

  - Example: The level sets for the function $f(x, y) = x^2 + y^2$ are circles in the plane. More specifically, the level set $x^2 + y^2 = c$ (for $c > 0$) is a circle with radius $\sqrt{c}$ centered at $(0, 0)$. For $c = 0$ the level set is just the single point $(0, 0)$, and for $c < 0$ the level sets do not contain any points at all. The first graph is the level set $x^2 + y^2 = 1$; the second graph contains level sets for $c = 1, 2, 3, \cdots, 8, 9$.

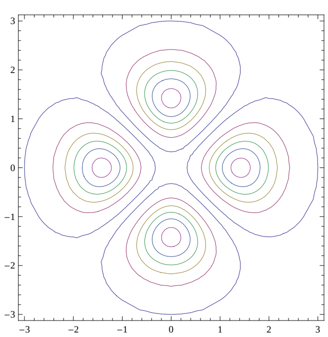
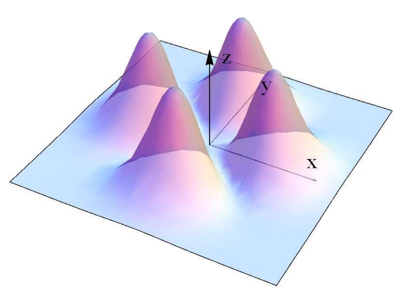



  - Example: Here are two collections of level curves for the function $f(x, y) = x^2 - y^2$, along with the 3D graph $z = f(x, y)$. The first set of level curves is $f(x, y) = c$ for $c = -4, -3, -2, -1, 0$, and the second set of level curves is $f(x, y) = c$ for $c = 0, 1, 2, 3, 4$.

    * The graph of $z = x^2 - y^2$ is called a hyperbolic paraboloid (or more colloquially, a "saddle") since it curves upward along the x-direction but downward along the y-direction. [The hyperbolic paraboloid is called that because it looks like a hyperbola in one cross-section, and a parabola in two others.]

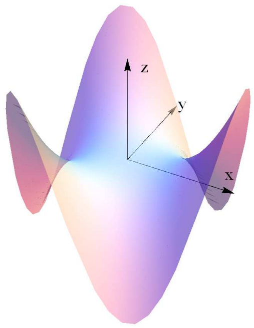
- Example: Here are some level curves for the function $f(x, y) = (x^2 - y^2)^2 e^{-x^2 - y^2}$, along with a 3D graph of $z = f(x, y)$, both plotted on the region $-3 \le x \le 3$ and $-3 \le y \le 3$:

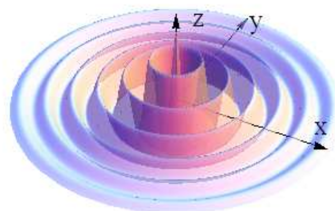



  - As one can see from comparing the plots, the level curves indicate where the function is changing in value: many level curves grouped closely together indicates that the function is changing quickly so that the function's graph will be "steep", while having level curves grouped far apart means that the function's graph will be fairly flat.

- Here are a few more graphs $z = f(x, y)$ for functions $f(x, y)$:

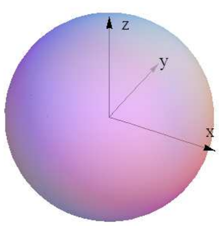
- Example: The graph $z = x^3 - 3xy^2$ is sometimes called the "monkey saddle", as it has three depressions rather than the two for the regular saddle (one for each leg, and one for the tail).
- Example: The graph $z = e^{3-\sqrt{x^2+y^2}/12} \cdot \cos\left(\sqrt{x^2+y^2}\right)$ produces a surface that looks like ripples in a pool of water.

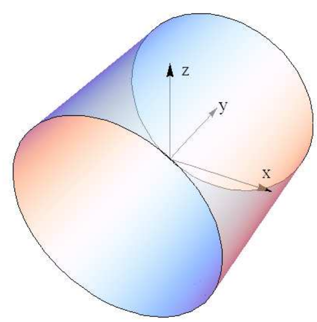




1.1.2 Graphing Functions of 3 Variables: Level Surfaces

- We can also represent some information graphically for functions $f(x, y, z)$ of three variables. Unfortunately, we cannot really produce proper graphs of these functions, since plotting a graph $w = f(x, y, z)$ would require drawing a 4-dimensional picture.
- However, we can still talk about level sets of functions of 3 variables – these are the points (x, y, z) satisfying a relation $f(x, y, z) = c$: these will (in general) give rise to level surfaces in 3-dimensional space.
  - Any graph $z = g(x, y)$ is an example of a level surface for the function $f(x, y, z) = g(x, y) - z$: the points (x, y, z) with $z = g(x, y)$ are the same as those with $f(x, y, z) = 0$.
- Example: The level surfaces for the function $f(x, y, z) = x^2 + y^2 + z^2$ are spheres centered at the origin: by the distance formula, the points satisfying $x^2 + y^2 + z^2 = c$ (for $c > 0$) are precisely those which are at a distance of $\sqrt{c}$ from the origin – but this is just another way of describing the sphere of radius $\sqrt{c}$ centered at $(0, 0, 0)$. A graph of the sphere $x^2 + y^2 + z^2 = 1$ is below.

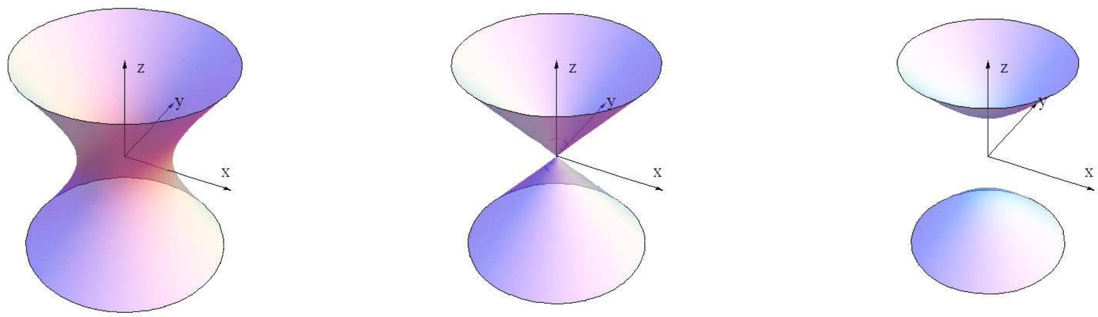




- Example: The level surfaces for the function $f(x, y, z) = x^2 + z^2$ are (right circular) cylinders running along the y-axis: again, from the distance formula, points satisfying $x^2 + z^2 = c$ (for $c > 0$) are those which are at a distance of $\sqrt{c}$ from the y-axis – but this is just another way of describing a right circular cylinder oriented along the y-axis. A graph of the cylinder $x^2 + z^2 = 1$ is above.
- Example: Let us examine the level surfaces for the function $f(x, y, z) = x^2 + y^2 - z^2$.
  - The set of points satisfying $x^2 + y^2 - z^2 = c$ for $c > 0$ forms a surface called a hyperboloid of one sheet, so named because the shape is a hyperbola in two cross-sections and an ellipse in the third, and the graph is connected ("one sheet").

* Remark: A hyperboloid of one sheet whose cross-sections are circles is an example of what is called a ruled surface: through each point on the surface pass two lines which are contained in the surface. Such surfaces can therefore be (physically) constructed using materials which are not curved. The hyperboloid of one sheet, in particular, is a common design for cooling towers.

- The set of points satisfying $x^2 + y^2 - z^2 = 0$ forms a (right circular) double cone whose axis is the z-axis and whose center point at the origin $(0, 0, 0)$.
- The set of points satisfying $x^2 + y^2 - z^2 = c$ for $c < 0$ forms a surface called a hyperboloid of two sheets, so named because the shape is a hyperbola in two cross-sections and an ellipse in the third, and the graph consists of two pieces ("two sheets").
- Here are plots of the level surfaces for $c = 1$, $c = 0$, and $c = -1$ respectively:



1.2 Partial Derivatives

- Partial derivatives are simply the usual notion of differentiation applied to functions of more than one variable. However, since we now have more than one variable, we also have more than one natural way to compute a derivative.
  - As with the definition of the derivative of a one-variable function, the derivatives of a function of several variables are formally defined using limits.
  - But like in the one-variable case, the formal definition of limit is cumbersome and generally not easy to use, even for simple functions.
  - Fortunately, we will not need to use limits to compute partial derivatives, since for all of the functions we will discuss, the computation of partial derivatives reduces to a version of the usual one-variable derivative.

- Definition: For a function $f(x, y)$ of two variables, we define the partial derivative of f with respect to x as $\dfrac{\partial f}{\partial x} = f_x = \lim_{h \to 0} \dfrac{f(x+h, y) - f(x, y)}{h}$ and the partial derivative of f with respect to y as $\dfrac{\partial f}{\partial y} = f_y = \lim_{h \to 0} \dfrac{f(x, y+h) - f(x, y)}{h}$.
  - Notation: In multivariable calculus, we use the symbol ∂ (typically pronounced either like the letter d or as "del") to denote taking a derivative, in contrast to single-variable calculus where we use the symbol d.
  - We will frequently use both notations $\dfrac{\partial f}{\partial y}$ and f_y to denote partial derivatives: we generally use the difference quotient notation when we want to emphasize a formal property of a derivative, and the subscript notation when we want to save space.
  - Geometrically, the partial derivative f_x captures how fast the function f is changing in the x-direction, and f_y captures how fast f is changing in the y-direction.

- To evaluate a partial derivative of the function f with respect to x, we need only pretend that all the other variables (i.e., everything except x) that f depends on are constants, and then just evaluate the derivative of f with respect to x as a normal one-variable derivative.

- All of the derivative rules (the Product Rule, Quotient Rule, Chain Rule, etc.) from one-variable calculus still hold: there will just be extra variables floating around.

- Example: Find f_x and f_y for $f(x, y) = x^3 y^2 + e^x$.
  - For f_x, we treat y as a constant and x as the variable. Thus, we see that $\boxed{f_x = 3x^2 \cdot y^2 + e^x}$.
  - Similarly, to find f_y, we instead treat x as a constant and y as the variable, to get $\boxed{f_y = x^3 \cdot 2y + 0 = 2x^3 y}$. (Note in particular that the derivative of e^x with respect to y is zero.)

- Example: Find f_x and f_y for $f(x, y) = \ln(x^2 + y^2)$.
  - For f_x, we treat y as a constant and x as the variable. We can apply the Chain Rule to get $\boxed{f_x = \dfrac{2x}{x^2 + y^2}}$, since the derivative of the inner function $x^2 + y^2$ with respect to x is $2x$.
  - Similarly, we can use the Chain Rule to find the partial derivative $\boxed{f_y = \dfrac{2y}{x^2 + y^2}}$.

- Example: Find f_x and f_y for $f(x, y) = \dfrac{e^{x\,y}}{x^2 + x}$.
  - For f_x we apply the Quotient Rule: $f_x = \dfrac{\frac{\partial}{\partial x}\left[e^{xy}\right] \cdot (x^2 + x) - e^{xy} \cdot \frac{\partial}{\partial x}\left[x^2 + x\right]}{(x^2 + x)^2}$. Then we can evaluate the derivatives in the numerator to get $\boxed{f_x = \dfrac{(y\, e^{xy}) \cdot (x^2 + x) - e^{xy} \cdot (2x + 1)}{(x^2 + x)^2}}$.
  - For f_y, the calculation is easier because the denominator is not a function of y. So in this case, we just need to use the Chain Rule to see that $\boxed{f_y = \dfrac{1}{x^2 + x} \cdot (x\, e^{xy})}$.

- We can generalize partial derivatives to functions of more than two variables: for each input variable, we get a partial derivative with respect to that variable. The procedure remains the same: treat all variables except the variable of interest as constants, and then differentiate with respect to the variable of interest.

- Example: Find f_x, f_y, and f_z for $f(x, y, z) = y\, z\, e^{2x^2 - y}$.
  - By the Chain Rule we have $\boxed{f_x = y\, z \cdot e^{2x^2 - y} \cdot 4x}$. (We don't need the Product Rule for f_x since y and z are constants.)
  - For f_y we need to use the Product Rule since f is a product of two nonconstant functions of y. We get $f_y = z \cdot e^{2x^2 - y} + y\, z \cdot \frac{\partial}{\partial y}\left[e^{2x^2 - y}\right]$, and then using the Chain Rule gives $\boxed{f_y = z\, e^{2x^2 - y} - y\, z \cdot e^{2x^2 - y}}$.
  - For f_z, all of the terms except for z are constants, so we have $\boxed{f_z = y\, e^{2x^2 - y}}$.

- Like in the one-variable case, we also have higher-order partial derivatives, obtained by taking a partial derivative of a partial derivative.
  - For a function of two variables, there are four second-order partial derivatives $f_{xx} = \frac{\partial}{\partial x}[f_x]$, $f_{xy} = \frac{\partial}{\partial y}[f_x]$, $f_{yx} = \frac{\partial}{\partial x}[f_y]$, and $f_{yy} = \frac{\partial}{\partial y}[f_y]$.
  - Remark: Partial derivatives in subscript notation are applied left-to-right, while partial derivatives in differential operator notation are applied right-to-left. (In practice, the order of the partial derivatives rarely matters, as we will see.)

- Example: Find the second-order partial derivatives f_{xx}, f_{xy}, f_{yx}, and f_{yy} for $f(x, y) = x^3 y^4 + y\, e^{2x}$.
  - First, we have $f_x = 3x^2 y^4 + 2y\, e^{2x}$ and $f_y = 4x^3 y^3 + e^{2x}$.

- Then we have $f_{xx} = \frac{\partial}{\partial x}\left[3x^2y^4 + 2y\,e^{2x}\right] = \boxed{6xy^4 + 4y\,e^{2x}}$ and $f_{xy} = \frac{\partial}{\partial y}\left[3x^2y^4 + 2y\,e^{2x}\right] = \boxed{12x^2y^3 + 2e^{2x}}$.
- Also we have $f_{yx} = \frac{\partial}{\partial x}\left[4x^3y^3 + e^{2x}\right] = \boxed{12x^2y^3 + 2e^{2x}}$ and $f_{yy} = \frac{\partial}{\partial y}\left[4x^3y^3 + e^{2x}\right] = \boxed{12x^3y^2}$.

- Notice that $f_{xy} = f_{yx}$ for the function in the example above. This is not an accident:
- Theorem (Clairaut): If both partial derivatives f_{xy} and f_{yx} are continuous, then they are equal.
  - In other words, these "mixed partials" are always equal (given mild assumptions about continuity), so there are really only three second-order partial derivatives.
  - This theorem can be proven using the limit definition of derivative and the Mean Value Theorem, but the details are unenlightening.
- We can continue on and take higher-order partial derivatives. For example, a function $f(x, y)$ has eight third-order partial derivatives: f_{xxx}, f_{xxy}, f_{xyx}, f_{xyy}, f_{yxx}, f_{yxy}, f_{yyx}, and f_{yyy}.
  - By Clairaut's Theorem, we can reorder the partial derivatives any way we want (if they are continuous, which is almost always the case). Thus, $f_{xxy} = f_{xyx} = f_{yxx}$, and $f_{xyy} = f_{yxy} = f_{yyx}$.
  - So in fact, $f(x, y)$ only has four different third-order partial derivatives: f_{xxx}, f_{xxy}, f_{xyy}, f_{yyy}.
- Example: Find the third-order partial derivatives f_{xxx}, f_{xxy}, f_{xyy}, f_{yyy} for $f(x, y) = x^4y^2 + x^3e^y$.
  - First, we have $f_x = 4x^3y^2 + 3x^2e^y$ and $f_y = 2x^4y + x^3e^y$.
  - Next, $f_{xx} = 12x^2y^2 + 6xe^y$, $f_{xy} = 8x^3y + 3x^2e^y$, and $f_{yy} = 2x^4 + x^3e^y$.
  - Finally, $f_{xxx} = \boxed{24xy^2 + 6e^y}$, $f_{xxy} = \boxed{24x^2y + 6xe^y}$, $f_{xyy} = \boxed{8x^3 + 3x^2e^y}$, and $f_{yyy} = \boxed{x^3e^y}$.
- Example: If all 5th-order partial derivatives of $f(x, y, z)$ are continuous and $f_{xyz} = e^{xyz}$, what is f_{zzyyx}?
  - By Clairaut's theorem, we can differentiate in any order, and so $f_{zzyyx} = f_{xyzyz} = (f_{xyz})_{yz}$.
  - Since $f_{xyz} = e^{xyz}$ we obtain $(f_{xyz})_y = xze^{xyz}$ and then $(f_{xyz})_{yz} = \boxed{xe^{xyz} + x^2yze^{xyz}}$.

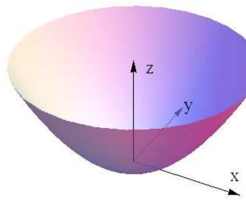
1.3 Local Extreme Points and Optimization

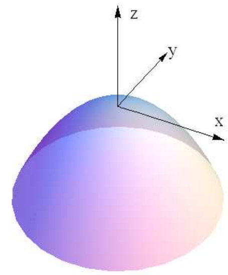
- Now that we have developed the basic ideas of derivatives for functions of several variables, we would like to know how to find minima and maxima of functions of several variables.
- We will primarily discuss functions of two variables, because there is a not-too-hard criterion for deciding whether a critical point is a minimum or a maximum.
  - Classifying critical points for functions of more than two variables requires some results from linear algebra, so we will not treat functions of more than two variables except for a special case we discuss in the next section.

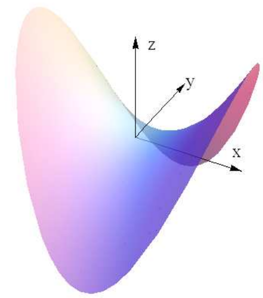
1.3.1 Minima, Maxima, and Saddle Points

- Definition: A local minimum is a critical point where f is "nearby" always bigger, a local maximum is a critical point where f is "nearby" always smaller, and a saddle point is a critical point where f "nearby" is bigger in some directions and smaller in others.
  - Example: The function $g(x, y) = x^2 + y^2$ has a local minimum at the origin.
  - Example: The function $p(x, y) = -(x^2 + y^2)$ has a local maximum at the origin.
  - Example: The function $h(x, y) = x^2 - y^2$ has a saddle point at the origin, since $h > 0$ along the x-axis (since $h(x, 0) = x^2$) but $h < 0$ along the y-axis (since $h(0, y) = -y^2$).

- Here are plots of the three examples:







- We would first like to determine where a function f can have a minimum or maximum value.
  - If $f_x(P) > 0$, then by moving slightly in the positive x-direction the value of f will increase, and by moving slightly in the negative x-direction the value of f will decrease.
  - Inversely, if $f_x(P) < 0$, then by moving slightly in the negative x-direction the value of f will increase, and by moving slightly in the positive x-direction the value of f will decrease.
  - Thus, f can only have a local minimum or maximum at P if $f_x(P) = 0$. By the same reasoning with y in place of x, we must also have $f_y(P) = 0$ at a local minimum or maximum.
- Definition: A critical point of the function $f(x, y)$ is a point (x_0, y_0) such that $f_x(x_0, y_0) = 0 = f_y(x_0, y_0)$, or either $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ is undefined.
  - By the observations above, a local minimum or maximum of a function can only occur at a critical point.
- Example: Find the critical points of the function $g(x, y) = x^2 + y^2 + 3$.
  - We have $g_x = 2x$ and $g_y = 2y$. Since both partial derivatives are defined everywhere, the only critical points will occur when $g_x = g_y = 0$.
  - We see $g_x = 0$ precisely when $x = 0$ and $g_y = 0$ precisely when $y = 0$.
  - Thus, there is a unique critical point $(x, y) = \boxed{(0, 0)}$.
- Example: Find the critical points of the function $h(x, y) = x^2 - 4x + y^2 + 2y$.
  - We have $h_x = 2x - 4$ and $h_y = 2y + 2$. Since both partial derivatives are defined everywhere, the only critical points will occur when $h_x = h_y = 0$.
  - We see $h_x = 0$ precisely when $x = 2$ and $h_y = 0$ precisely when $y = -1$.
  - Thus, there is a unique critical point $(x, y) = \boxed{(2, -1)}$.
- Example: Find the critical points of the function $q(x, y) = x^2 + 3xy + 2y^2 - 5x - 8y + 4$.
  - We have $q_x = 2x + 3y - 5$ and $q_y = 3x + 4y - 8$. Since both partial derivatives are defined everywhere, the only critical points will occur when $q_x = q_y = 0$.
  - This yields the equations $2x + 3y - 5 = 0$ and $3x + 4y - 8 = 0$.
  - Setting $2x + 3y - 5 = 0$ and solving for y yields $y = (5 - 2x)/3$. Plugging into the second equation and simplifying eventually gives $x/3 - 4/3 = 0$, so that $x = 4$ and then $y = -1$.
  - Therefore, there is a single critical point $(x, y) = \boxed{(4, -1)}$.
- Example: Find the critical points of the function $p(x, y) = x^3 + y^3 - 3xy$.
  - We have $p_x = 3x^2 - 3y$ and $p_y = 3y^2 - 3x$. Since both partial derivatives are defined everywhere, the only critical points will occur when $p_x = p_y = 0$.

- This gives the two equations $3x^2 - 3y = 0$ and $3y^2 - 3x = 0$, or, equivalently, $x^2 = y$ and $y^2 = x$.
- Plugging the first equation into the second yields $x^4 = x$: thus, $x^4 - x = 0$, and factoring yields $x(x-1)(x^2+x+1) = 0$.
- The only real solutions are $x = 0$ (which then gives $y = x^2 = 0$) and $x = 1$ (which then gives $y = x^2 = 1$).
- Therefore, there are two critical points: $\boxed{(0,0) \text{ and } (1,1)}$.

- Example: Find the critical points of the function $f(x,y) = x\, e^{y^2-2x^2}$.
  - We have $f_x = e^{y^2-2x^2} + xe^{y^2-2x^2} \cdot (-4x) = (1-4x^2)e^{y^2-2x^2}$ and $f_y = x\, e^{y^2-2x^2} \cdot (2y) = 2xy\, e^{y^2-2x^2}$. Since both partial derivatives are defined everywhere, the only critical points will occur when $f_x = f_y = 0$.
  - Since exponentials are never zero, we see that $f_x = 0$ when $1 - 4x^2 = 0$ so that $x = \pm\frac{1}{2}$, while $f_y = 0$ when $2xy = 0$ so that $x = 0$ or $y = 0$. Since x cannot be zero by the first equation (since $x = \pm\frac{1}{2}$) we must have $y = 0$.
  - Therefore, there are two critical points: $\boxed{(\frac{1}{2}, 0) \text{ and } (-\frac{1}{2}, 0)}$.

1.3.2 Classifying Critical Points

- Now that we have a list of critical points (namely, the places that a function could potentially have a minimum or maximum value) we would like to know whether those points actually are minima or maxima of f.
- Definition: The discriminant (also called the Hessian) at a critical point is the value $D = f_{xx} \cdot f_{yy} - (f_{xy})^2$, where each of the second-order partials is evaluated at the critical point.
  - One way to remember the definition of the discriminant is as the determinant of the matrix of the four second-order partials: $D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$. (We are implicitly using the fact that $f_{xy} = f_{yx}$.)
  - Example: For $g(x,y) = x^2 + y^2$ we have $g_{xx} = g_{yy} = 2$ and $g_{xy} = 0$ so $D = 4$ at the origin.
  - Example: For $h(x,y) = x^2 - y^2$ we have $h_{xx} = 2$, $h_{yy} = -2$, and $h_{xy} = 0$ so $D = -4$ at the origin.
  - Remark: The reason this value is named "discriminant" can be seen by computing D for the function $p(x,y) = ax^2 + bxy + cy^2$: the result is $D = 4ac - b^2$, which is -1 times the quantity $b^2 - 4ac$, the famous discriminant for the quadratic polynomial $ax^2 + bx + c$. (Recall that the discriminant of $ax^2 + bx + c$ determines how many real roots the polynomial has.)
- Theorem (Second Derivatives Test): Suppose P is a critical point of $f(x,y)$, and let D be the value of the discriminant $f_{xx}f_{yy} - f_{xy}^2$ at P. If $D > 0$ and $f_{xx} > 0$, then the critical point is a minimum. If $D > 0$ and $f_{xx} < 0$, then the critical point is a maximum. If $D < 0$, then the critical point is a saddle point. (If $D = 0$, then the test is inconclusive.)
  - Proof (outline): Assume for simplicity that P is at the origin. Then one may show that the function $f(x,y) - f(P)$ is closely approximated by the polynomial $ax^2 + bxy + cy^2$, where $a = \frac{1}{2}f_{xx}$, $b = f_{xy}$, and $c = \frac{1}{2}f_{yy}$. If $D \neq 0$, then the behavior of $f(x,y)$ near the critical point P will be the same as that quadratic polynomial. Completing the square and examining whether the resulting quadratic polynomial has any real roots and whether it opens or downwards yields the test.
- We can combine the above results to yield a procedure for finding and classifying the critical points of a function $f(x,y)$:
  - Step 1: Compute both partial derivatives f_x and f_y.
  - Step 2: Find all points (x,y) where *both* partial derivatives are zero, or where (at least) *one* of the partial derivatives is undefined.

    * It may require some algebraic manipulation to find the solutions: a basic technique is to solve one equation for one of the variables, and then plug the result into the other equation. Another technique is to try to factor one of the equations and then analyze cases.

- Step 3: At *each* critical point, evaluate $D = f_{xx} \cdot f_{yy} - (f_{xy})^2$ and apply the Second Derivatives Test:

If $\boxed{D > 0 \text{ and } f_{xx} > 0\text{: local minimum}}$. If $\boxed{D > 0 \text{ and } f_{xx} < 0\text{: local maximum}}$. If $\boxed{D < 0\text{: saddle point}}$.

- Example: Verify that $f(x, y) = x^2 + y^2$ has only one critical point, a minimum at the origin.
    - First, we have $f_x = 2x$ and $f_y = 2y$. Since they are both defined everywhere, we need only find where they are both zero.
    - Setting both partial derivatives equal to zero yields $x = 0$ and $y = 0$, so the only critical point is $(0, 0)$.
    - To classify the critical points, we compute $f_{xx} = 2$, $f_{xy} = 0$, and $f_{yy} = 2$. Then $D = 2 \cdot 2 - 0^2 = 4$.
    - So, by the classification test, since $D > 0$ and $f_{xx} > 0$ at $(0, 0)$, we see that $\boxed{(0, 0) \text{ is a local minimum}}$.

- Example: For the function $f(x, y) = 3x^2 + 2y^3 - 6xy$, find the critical points and classify them as minima, maxima, or saddle points.
    - First, we have $f_x = 6x - 6y$ and $f_y = 6y^2 - 6x$. Since they are both defined everywhere, we need only find where they are both zero.
    - Next, we can see that f_x is zero only when $y = x$. Then the equation $f_y = 0$ becomes $6x^2 - 6x = 0$, which by factoring we can see has solutions $x = 0$ or $x = 1$. Since $y = x$, we conclude that $\boxed{(0, 0), \text{ and } (1, 1) \text{ are critical points}}$.
    - To classify the critical points, we compute $f_{xx} = 6$, $f_{xy} = -6$, and $f_{yy} = 12y$. Then $D(0, 0) = 6 \cdot 0 - (-6)^2 < 0$ and $D(1, 1) = 6 \cdot 12 - (-6)^2 > 0$.
    - So, by the classification test, $\boxed{(0, 0) \text{ is a saddle point}}$ and $\boxed{(1, 1) \text{ is a local minimum}}$.

- Example: For the function $g(x, y) = x^3y - 3xy^3 + 8y$, find the critical points and classify them as minima, maxima, or saddle points.
    - First, we have $g_x = 3x^2y - 3y^3$ and $g_y = x^3 - 9xy^2 + 8$. Since they are both defined everywhere, we need only find where they are both zero.
    - Setting both partial derivatives equal to zero. Since $g_x = 3y(x^2 - y^2) = 3y(x + y)(x - y)$, we see that $g_x = 0$ precisely when $y = 0$ or $y = x$ or $y = -x$.
    - If $y = 0$, then $g_y = 0$ implies $x^3 + 8 = 0$, so that $x = -2$. This yields the point $(x, y) = (-2, 0)$.
    - If $y = x$, then $g_y = 0$ implies $-8x^3 + 8 = 0$, so that $x = 1$. This yields the point $(x, y) = (1, 1)$.
    - If $y = -x$, then $g_y = 0$ implies $-8x^3 + 8 = 0$, so that $x = 1$. This yields the point $(x, y) = (1, -1)$.
    - To summarize, we see that $\boxed{(-2, 0), (1, 1), \text{ and } (1, -1) \text{ are critical points}}$.
    - To classify the critical points, we compute $g_{xx} = 6xy$, $g_{xy} = 3x^2 - 9y^2$, and $g_{yy} = -18xy$.
    - Then $D(-2, 0) = 0 \cdot 0 - (12)^2 < 0$, $D(1, 1) = 6 \cdot (-18) - (-6)^2 < 0$, and $D(1, -1) = (-6) \cdot (18) - (-6)^2 < 0$.
    - So, by the classification test, $\boxed{(-2, 0), (1, 1), \text{ and } (1, -1) \text{ are all saddle points}}$.

- Example: Find the value of the function $h(x, y) = x + 2y^4 - \ln(x^4y^8)$ at its local minimum, for x and y positive.
    - To solve this problem, we will search for all critical points of $h(x, y)$ that are minima.
    - First, we have $h_x = 1 - \dfrac{4x^3y^8}{x^4y^8} = 1 - \dfrac{4}{x}$ and $h_y = 8y^3 - \dfrac{8x^4y^7}{x^4y^8} = 8y^3 - \dfrac{8}{y}$. Both partial derivatives are defined everywhere in the given domain.

- We see that $h_x = 0$ only when $x = 4$, and also that $h_y = 0$ is equivalent to $\frac{8}{y}(y^4 - 1) = 0$, which holds for $y = \pm 1$. Since we only want $y > 0$, there is a unique critical point: $(4, 1)$.
- Next, we compute $h_{xx} = \frac{4}{x^2}$, $g_{xy} = 0$, and $g_{yy} = 24y^2 + \frac{8}{y^2}$. Then $D(4, 1) = \frac{1}{4} \cdot 32 - 0^2 > 0$.
- Thus, there is a unique critical point, and it is a minimum. Therefore, we conclude that the function has a local minimum at $(4, 1)$, and the minimum value is $h(4, 1) = \boxed{6 - \ln(4^4)}$.

- Example: Find the minimum distance between a point on the plane $x + y + z = 1$ and the point $(2, -1, -2)$.
  - The distance from the point (x, y, z) to $(2, -1, 2)$ is $d = \sqrt{(x-2)^2 + (y+1)^2 + (z+2)^2}$. Since $x+y+z = 1$ on the plane, we can view this as a function of x and y only: $d(x, y) = \sqrt{(x-2)^2 + (y+1)^2 + (3-x-y)^2}$.
  - We could minimize $d(x, y)$ by finding its critical points and searching for a minimum, but it will be much easier to find the minimum value of the squared distance $f(x, y) = d(x, y)^2 = (x-2)^2+(y+1)^2+(3-x-y)^2$.
  - We compute $f_x = 2(x-2) - 2(3-x-y) = 4x + 2y - 10$ and $f_y = 2(y+1) - 2(3-x-y) = 2x + 4y - 4$. Both partial derivatives are defined everywhere, so we need only find where they are both zero.
  - Setting $f_x = 0$ and solving for y yields $y = 5 - 2x$, and then plugging this into $f_y = 0$ yields $2x + 4(5 - 2x) - 4 = 0$, so that $-6x + 16 = 0$. Thus, $x = 8/3$ and then $y = -1/3$.
  - Furthermore, we have $f_{xx} = 4$, $f_{xy} = 2$, and $f_{yy} = 4$, so that $D = f_{xx}f_{yy} - f_{xy}^2 = 12 > 0$. Thus, the point $(x, y) = (8/3, -1/3)$ is a local minimum.
  - Thus, there is a unique critical point, and it is a minimum. We conclude that the distance function has its minimum at $(4, 1)$, so the minimum distance is $d(8/3, -1/3) = \sqrt{(2/3)^2 + (2/3)^2 + (2/3)^2} = \boxed{2/\sqrt{3}}$.

1.4 Optimization of a Function on a Region, Linear Programming

- We now discuss the problem of finding the minimum and maximum values of a function on a region in the plane, rather than the entire plane itself.
  - In general, if the region is not closed (i.e., does not contain its boundary, like the region $x^2 + y^2 < 1$ which does not contain the boundary circle $x^2 + y^2 = 1$) or not bounded (i.e., extends infinitely far away from the origin, like the half-plane $x \geq 0$) then a continuous function may not attain its minimum or maximum values anywhere in the region.
  - In order to ensure that a function does attain its minimum and maximum values at some point inside the region, the region must be both closed and bounded. If the region is not bounded or not closed, we must additionally study what happens to the function as we approach the region's boundary, or what happens as we move far away from the origin.

1.4.1 Optimization on a Region

- A natural first step is to find the critical points of the function. However, if we want to find the absolute minimum or maximum of a function $f(x, y)$ on a closed and bounded region, we must also analyze the function's behavior on the boundary of the region, because the boundary could contain the minimum or maximum.
  - Example: The extreme values of $f(x, y) = x^2 - y^2$ on the square $0 \leq x \leq 1$, $0 \leq y \leq 1$ occur at two of the "corner points": the minimum is -1 occurring at $(0, 1)$, and the maximum $+1$ occurring at $(1, 0)$. We can see that these two points are actually the minimum and maximum on this region without calculus: since squares of real numbers are always nonnegative, on the region in question we have $-1 \leq -y^2 \leq x^2 - y^2 \leq x^2 \leq 1$.

- Unfortunately, unlike the case of a function of one variable where the boundary of an interval $[a, b]$ is very simple (namely, the two values $x = a$ and $x = b$), the boundary of a region in the plane or in higher-dimensional space can be rather complicated.

- Ultimately, one needs to find a parametrization $(x(t), y(t))$ of the boundary of the region, or some other description. (This may require breaking the boundary into several pieces, depending on the shape of the region.)
- Then, by plugging the parametrization of the boundary curve into the function, we obtain a function $f(x(t),\, y(t))$ of the single variable t, which we can then analyze to determine the behavior of the function on the boundary.

- To find the absolute minimum and maximum values of a function on a given closed and bounded region R, follow these steps:
  - Step 1: Find all of the critical points of f that lie inside the region R.
  - Step 2: Parametrize the boundary of the region R (separating into several components if necessary) as $x = x(t)$ and $y = y(t)$, then plug in the parametrization to obtain a function of t, $f(x(t), y(t))$. Then search for "boundary-critical" points, where the t-derivative $\frac{d}{dt}$ of $f(x(t), y(t))$ is zero. Also include endpoints, if the boundary components have them.
    - A line segment from (a, b) to (c, d) can be parametrized by $x(t) = a + t(c - a)$, $y(t) = b + t(d - b)$, for $0 \le t \le 1$.
    - A curve of the form $y = g(x)$ can be parametrized by $x(t) = t$, $y(t) = g(t)$.
  - Step 3: Plug the full list of critical and boundary-critical points into f, and find the largest and smallest values.

- Example: Find the absolute maximum and minimum of $f(x, y) = x^2 - xy + y$ on the rectangle $0 \le x \le 2$, $0 \le y \le 3$.
  - First, we find the critical points: since $f_x = 2x - y$ and $f_y = -x + 1$, there is a single critical point $\boxed{(1, 2)}$.
  - Next, we analyze the boundary of the region, which has 4 components:
    - Component #1, a line segment from $(0, 0)$ to $(2, 0)$: This component is parametrized by $x = 2t$, $y = 0$ for $0 \le t \le 1$. On this component we have $f(2t, 0) = 4t^2$, which has a critical point at $t = 0$ corresponding to $(x, y) = \boxed{(0, 0)}$. We also have boundary points $\boxed{(0, 0),\ (2, 0)}$.
    - Component #2, a line segment from $(2, 0)$ to $(2, 3)$: This component is parametrized by $x = 2$, $y = 3t$ for $0 \le t \le 1$. On this component we have $f(2, 3t) = 4 - 3t$, which has no critical point. We get only the boundary points $\boxed{(2, 0),\ (2, 3)}$.
    - Component #3, a line segment from $(0, 3)$ to $(2, 3)$: This component is parametrized by $x = 2t$, $y = 3$ for $0 \le t \le 1$. On this component we have $f(2t, 3) = 4t^2 - 6t + 3$, which has derivative $8t - 6$ hence has a critical point at $t = 3/4$ corresponding to $(x, y) = \boxed{(3/2, 3)}$. We also have boundary points $\boxed{(0, 3),\ (2, 3)}$.
    - Component #4, a line segment from $(0, 0)$ to $(0, 3)$: This component is parametrized by $x = 0$, $y = 3t$ for $0 \le t \le 1$. On this component we have $f(0, 3t) = 3$, which has no critical point. We get only the boundary points $\boxed{(0, 0),\ (0, 3)}$.
  - Our full list of points to analyze is $(1, 2)$, $(0, 0)$, $(2, 0)$, $(2, 3)$, $(3/2, 3)$, $(0, 3)$. We have $f(1, 2) = 1$, $f(0, 0) = 0$, $f(2, 0) = 4$, $f(2, 3) = 1$, $f(3/2, 3) = 3/4$, and $f(0, 3) = 3$. The $\boxed{\text{maximum is } 4}$ and the $\boxed{\text{minimum is } 0}$.

- Example: Find the absolute minimum and maximum of $f(x, y) = x^3 + 6xy - y^3$ on the triangle with vertices $(0, 0)$, $(4, 0)$, and $(0, -4)$.
  - First, we find the critical points: we have $f_x = 3x^2 + 6y$ and $f_y = -3y^2 + 6x$. Solving $f_y = 0$ yields $x = y^2/2$ and then plugging into $f_x = 0$ gives $y^4/4 + 2y = 0$ so that $y(y^3 + 8) = 0$: thus, we see that $\boxed{(0, 0) \text{ and } (2, -2) \text{ are critical points}}$.
  - Next, we analyze the boundary of the region. Here, the boundary has 3 components.

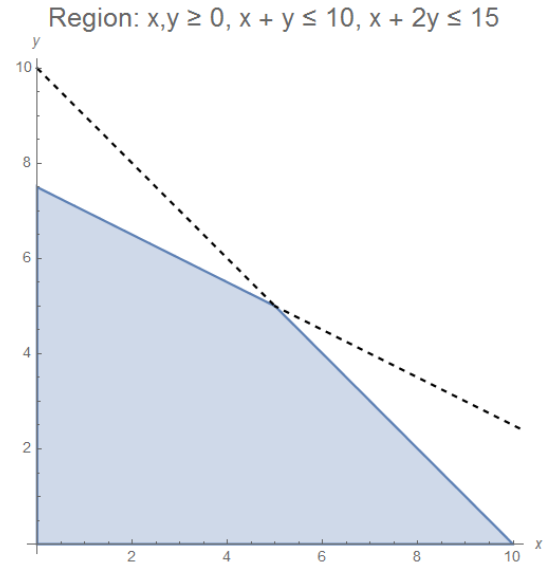
    * Component #1, joining $(0,0)$ to $(4,0)$: This component is parametrized by $x = t$, $y = 0$ for $0 \le t \le 4$. On this component we have $f(t,0) = t^3$, which has a critical point only at $t = 0$, which corresponds to $(x,y) = \boxed{(0,0)}$. Also add the boundary point $\boxed{(4,0)}$.
    * Component #2, joining $(0,-4)$ to $(4,0)$: This component is parametrized by $x = t$, $y = t-4$ for $0 \le t \le 4$. On this component we have $f(t,t-4) = 18t^2 - 72t + 64$, which has a critical point for $t = 2$, corresponding to $(x,y) = \boxed{(2,-2)}$. Also add the boundary points $\boxed{(4,0)}$ and $\boxed{(0,-4)}$.
    * Component #3, joining $(0,0)$ to $(0,-4)$: This component is parametrized by $x = 0$, $y = -t$ for $0 \le t \le 4$. On this component we have $f(0,t) = t^3$, which has a critical point for $t = 0$, corresponding to $(x,y) = \boxed{(0,0)}$. Also add the boundary point $\boxed{(0,-4)}$.
  - Our full list of points to analyze is $(0,0)$, $(4,0)$, $(0,-4)$, and $(2,-2)$. We compute $f(0,0) = 0$, $f(4,0) = 64$, $f(0,-4) = 64$, $f(2,-2) = -8$, and so we see that $\boxed{\text{maximum is 64}}$ and the $\boxed{\text{minimum is -8}}$.

- Example: Find the absolute maximum and minimum of $f(x,y) = xy - 3x$ on the region with $x^2 \le y \le 9$.
  - First, we find the critical points: since $f_x = y - 3$ and $f_y = x$, there is a single critical point $\boxed{(0,3)}$.
  - Next, we analyze the boundary of the region, which (as a quick sketch reveals) has 2 components.
    * Component #1, a line segment from $(-3,9)$ to $(3,9)$: This component is parametrized by $x = t$, $y = 9$ for $-3 \le t \le 3$. On this component we have $f(t,9) = 6t$, which has no critical point. We only have boundary points $\boxed{(-3,9),\ (3,9)}$.
    * Component #2, a parabolic arc parametrized by $x = t$, $y = t^2$ for $-3 \le t \le 3$. On this component we have $f(t,t^2) = t^3 - 3t$, which has critical points at $t = \pm 1$, corresponding to $(x,y) = \boxed{(-1,1),\ (1,1)}$. The boundary points $(-3,9)$, $(3,9)$ are already listed above.
  - Our full list of points to analyze is $(0,3)$, $(-3,9)$, $(3,9)$, $(-1,1)$, $(1,1)$. We have $f(0,3) = 0$, $f(-3,9) = -18$, $f(3,9) = 18$, $f(-1,1) = 2$, and $f(1,1) = -2$. The $\boxed{\text{maximum is 18}}$ and the $\boxed{\text{minimum is -18}}$.

1.4.2 Linear Programming

- A particular special class of optimization problems involves searching for the minimum or maximum value of a linear function subject to various linear constraints (i.e., on a region defined by linear inequalities such as $3x + 2y \le 8$ or $x + z \ge 0$): these are known as linear programming problems[1].
- We can use the same methods for optimization of a function on a region to solve linear programming problems. However, linear programming problems have some convenient features that make it easier to identify potential minima and maxima, which we will illustrate with an example:
- Example: Find the minimum and maximum values of the function $f(x,y) = 2 + 3x + 5y$ subject to the constraints $x \ge 0$, $y \ge 0$, $x + y \le 10$, $x + 2y \le 15$.
  - First we search for critical points of f: since $f_x = 3$ and $f_y = 5$, there are no critical points.
  - Next, we need to determine the structure of the region, which is shown below:

[1] Although there are numerous computational algorithms (such as the simplex method) that exist for solving large-scale linear programming problems, the word "programming" in "linear programming" does not refer to computer programs. Instead, it comes from the United States military usage of the word "program" in reference to training and logistics schedules, whose optimization was among the first applied examples of linear programming.



- We can see that the boundary has 4 components, and we can find the intersection points of the various lines that make up components of the boundary by solving the constraint equations.
- Explicitly, the lines $x = 0$ and $y = 0$ intersect at the origin $(0, 0)$, the line $y = 0$ intersects $x + y = 10$ at $(10, 0)$, the line $x + y = 10$ intersects $x + 2y = 15$ at $(5, 5)$, and the line $x + 2y = 15$ intersects $x = 0$ at $(0, 15/2)$.
- Now we can parametrize the four components of the boundary:
  * Component #1, a line segment from $(0, 0)$ to $(10, 0)$: This component is parametrized by $x = 10t$, $y = 0$ for $0 \le t \le 1$. On this component we have $f(10t, 0) = 2 + 30t$, which has derivative 30 and thus has no critical points. We get only the boundary points $\boxed{(0,0),\ (10,0)}$.
  * Component #2, a line segment from $(5, 5)$ to $(10, 0)$: This component is parametrized by $x = 5+5t$, $y = 5-5t$ for $0 \le t \le 1$. On this component we have $f(5+5t, 5-5t) = 42-10t$, which has derivative -10 and thus has no critical point. We get only the boundary points $\boxed{(5,5),\ (10,0)}$.
  * Component #3, a line segment from $(0, 15/2)$ to $(5, 5)$: This component is parametrized by $x = 5t$, $y = \frac{15}{2} - \frac{5}{2}t$ for $0 \le t \le 1$. On this component we have $f(5t, \frac{15}{2} - \frac{5}{2}t) = \frac{79}{2} + \frac{5}{2}t$, which has derivative $\frac{5}{2}$ and thus has no critical point. We get only the boundary points $\boxed{(0,15/2),\ (5,5)}$.
  * Component #4, a line segment from $(0, 0)$ to $(0, 15/2)$: This component is parametrized by $x = 0$, $y = \frac{15}{2}t$ for $0 \le t \le 1$. On this component we have $f(0, \frac{15}{2}t) = 2 + \frac{75}{2}t$, which has derivative $\frac{75}{2}$ and thus has no critical point. We get only the boundary points $\boxed{(0,0),\ (0,15/2)}$.
- Our full list of points to analyze is $(0, 0)$, $(10, 0)$, $(5, 5)$, $(0, 15/2)$. We have $f(0, 0) = 2$, $f(10, 0) = 32$, $f(5, 5) = 42$, $f(0, 15/2) = 79/2$. We see that the $\boxed{\text{maximum is } 42}$ and the $\boxed{\text{minimum is } 2}$.

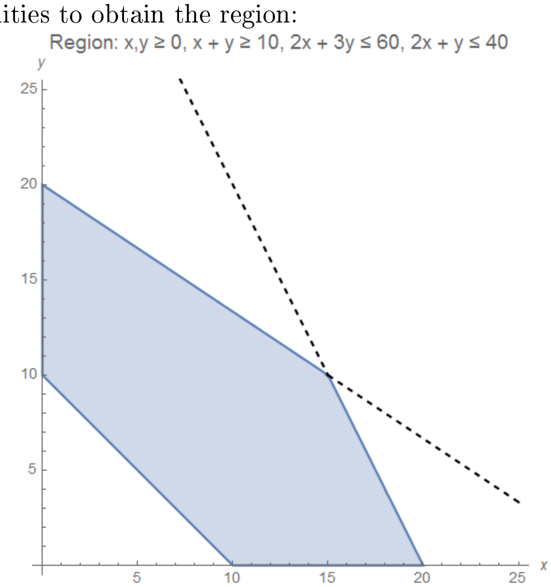
• In the example above, notice that we did not obtain any critical points nor any boundary-critical points: the only points on our candidate list were the "corner points" of the region.

- In fact, we can easily see that this will be the case for any linear programming problem (where the function to be optimized is linear, and the region is defined by linear inequalities).
- Since the function is linear, its partial derivatives are all constants, and so (unless all the partial derivatives are zero, in which case the function is constant) there will be no critical points.
- Likewise, since the boundary of the region can be parametrized using linear functions, when we evaluate the function on the boundary the resulting function will also be linear, and therefore its derivative will be constant. Then (unless the derivative is zero, in which case the function is constant) there will be no boundary-critical points.

- In all cases, we see that the minimum and maximum values will always be attained at one of the "corner points" of the boundary. (This observation is sometimes called the fundamental theorem of linear programming.)
- In fact, all of this analysis still holds for linear programming problems in more than 2 variables (although of course it is typically more difficult to identify all of the corner points by hand, unless the number of variables and inequalities defining the region is fairly small).

- To solve a linear programming problem, follow these steps:
  - Step 1: Identify the function f to be optimized as well as all of the constraint inequalities.
  - Step 2: Draw the region (if possible) and identify all of the corner points.
  - Step 3: Plug the full list of corner points into f, and find the largest and smallest values.
  - Note that in order for this procedure to apply, the region must be finite. If the region is infinite, it is also necessary to analyze the behavior of f on the unbounded portion of the region.

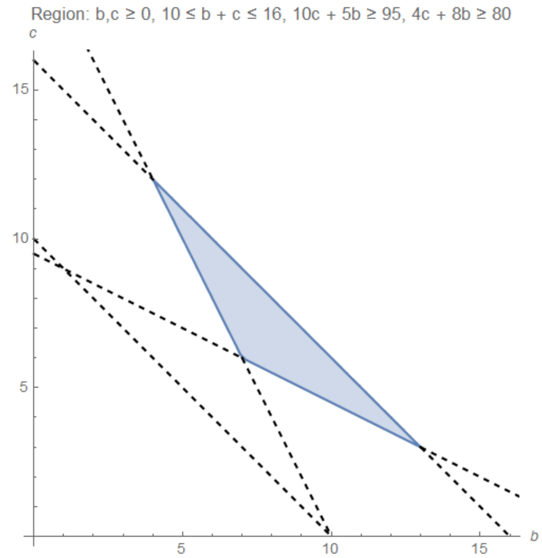
- Example: Find the minimum and maximum values of the function $P(x, y) = 4x + 3y + 5$ subject to the conditions $x \geq 0$, $y \geq 0$, $x + y \geq 10$, $2x + 3y \leq 60$, $2x + y \leq 50$.
  - We plot the various inequalities to obtain the region:



  - From the plot of the region, we can identify the corner points as follows:
    - Intersection of $x = 0$ with $x + y = 10$ and $2x + 3y = 60$, yielding the points $(0, 10)$ and $(0, 20)$.
    - Intersection of $y = 0$ with $x + y = 10$ and $2x + y = 40$, yielding the points $(10, 0)$ and $(20, 0)$.
    - Intersection of $2x + 3y = 60$ and $2x + y = 40$. Solving the second equation yields $y = 40 - 2x$, and plugging into the first equation gives $2x + 3(40 - 2x) = 60$ so that $-4x + 120 = 60$ so $x = 15$ and then $y = 10$, yielding the point $(15, 10)$.
  - Our list of candidate points is $\boxed{(0, 10),\ (0, 20),\ (10, 0),\ (20, 0),\ (15, 10)}$.
  - We compute $f(0, 10) = 35$, $f(0, 20) = 65$, $f(10, 0) = 45$, $f(20, 0) = 85$, and $f(15, 10) = 95$, so the $\boxed{\text{maximum is } 95}$ and the $\boxed{\text{minimum is } 35}$.

- Example: A cat-food company makes its food from chicken, which costs 25 cents per ounce, and beef, which costs 20 cents per ounce. Chicken has 10 grams of protein and 4 grams of fat per ounce, while beef has 5 grams of protein and 8 grams of fat per ounce. Each package of food must weigh between 10 and 16 ounces, and it must also have at least 95 grams of protein and at least 80 grams of fat. How much chicken and beef should the company use in each package to minimize the total cost while also satisfying these requirements?
  - Suppose that the company uses c ounces of chicken and b ounces of beef: then the total cost is $f(b, c) = 20b + 25c$ (which we wish to minimize).

- We also must satisfy the constraints $b \geq 0$ and $c \geq 0$ (since the cans cannot contain a negative amount of either ingredient), $10 \leq b + c \leq 16$ (for the weight), $10c + 5b \geq 95$ (for the protein), and $4c + 8b \geq 80$ (for the fat).
- The corresponding region is as follows:



- From the picture, we can see that the constraints $b \geq 0$, $c \geq 0$, and $b + c \geq 10$ are irrelevant and are not parts of the boundary of the region. There are three corner points, which we can find as follows:
  - Intersection of $b + c = 16$ with $10c + 5b = 95$. The first equation yields $c = 16 - b$, and plugging into the second equation yields $10(16 - b) + 5b = 95$ so that $160 - 5b = 95$. Thus $b = 13$ and then $c = 3$.
  - Intersection of $b + c = 16$ with $4c + 8b = 80$. The first equation yields $c = 16 - b$, and plugging into the second equation yields $4(16 - b) + 8b = 80$ so that $64 + 4b = 80$. Thus $b = 4$ and then $c = 12$.
  - Intersection of $10c + 5b = 95$ with $4c + 8b = 80$. The first equation yields $b = 19 - 2c$, and plugging into the second equation yields $4c + 8(19 - 2c) = 80$ so that $152 - 12c = 80$. Thus $c = 6$ and then $b = 7$.
- Thus we obtain three corner points, $(b, c) = (13, 3)$, $(4, 12)$, and $(7, 6)$.
- We compute $f(13, 3) = 335$, $f(4, 12) = 380$, and $f(7, 6) = 290$, so the minimum cost of \$2.90 occurs with $(b, c) = (7, 6)$, which is to say, with $\boxed{\text{7 ounces of beef and 6 ounces of chicken}}$.

1.5 Lagrange Multipliers and Constrained Optimization

- Many types of applied optimization problems are not of the form "given a function, maximize it on a region", but rather of the form "given a function, maximize it subject to some additional constraints".
  - <u>Example</u>: Maximize the volume $V = \pi r^2 h$ of a cylindrical can given that its surface area $SA = 2\pi r^2 + 2\pi r h$ is $150\pi\,\text{cm}^2$.
- The most natural way to attempt such a problem is to eliminate the constraints by solving for one of the variables in terms of the others and then reducing the problem to something without a constraint. Then we are able to perform the usual procedure of evaluating the derivative (or derivatives), setting them equal to zero, and looking among the resulting critical points for the desired extreme point.
  - In the example above, we would use the surface area constraint $150\pi\,\text{cm}^2 = 2\pi r^2 + 2\pi r h$ to solve for h in terms of r, obtaining $h = \dfrac{150\pi - 2\pi r^2}{2\pi r} = \dfrac{75 - r^2}{r}$, and then plug in to the volume formula to write it as a function of r alone: this gives $V(r) = \pi r^2 \cdot \dfrac{75 - r^2}{r} = 75\pi r - \pi r^3$.
  - Then $\dfrac{dV}{dr} = 75\pi - 3\pi r^2$, so setting equal to zero and solving shows that the critical points occur for $r = \pm 5$.

- Since we are interested in positive r, we can do a little bit more checking to conclude that the can's volume is indeed maximized at the critical point, so the radius is $r = 5\,\text{cm}$, the height is $h = 10\,\text{cm}$, and the resulting volume is $V = 250\pi\text{cm}^3$.

- Using the technique of Lagrange multipliers, however, we can perform a constrained optimization without having to solve the constraint equations. This technique is especially useful when the constraints are difficult or impossible to solve explicitly.

- Method (Lagrange multipliers, 1 constraint): To find the extreme values of $f(x,y,z)$ subject to a constraint $g(x,y,z) = c$, define the Lagrange function $L(x,y,z,\lambda) = f(x,y,z) - \lambda \cdot [g(x,y,z) - c]$. Then any extreme value of $f(x,y,z)$ subject to the constraint $g(x,y,z) = c$ must occur at a critical point of $L(x,y,z,\lambda)$. In other words, it is sufficient to solve the system of four variables x, y, z, λ given by $f_x = \lambda g_x$, $f_y = \lambda g_y$, $f_z = \lambda g_z$, $g(x,y,z) = c$, and then search among the resulting triples (x,y,z) to find the minimum and maximum.
  - If we have two variables, we would instead solve the system $f_x = \lambda g_x$, $f_y = \lambda g_y$, $g(x,y) = c$.
  - Remark: The value λ is called a Lagrange multiplier.
  - Here is the intuitive idea behind the method:
    - Imagine we are walking around the level set $g(x,y,z) = c$, and consider what the contours of $f(x,y,z)$ are doing as we move around.
    - In general the contours of f and g will be different, and they will cross one another.
    - But if we are at a point where f is maximized, then if we walk around nearby that maximum, we will see only contours of f with a smaller value than the maximum.
    - Thus, at that maximum, the contour $g(x,y,z) = c$ is tangent to the contour of f.
    - This information can, in turn, be reinterpreted as saying that the vector $\nabla f = \langle f_x, f_y, f_z\rangle$ is parallel to the vector $\nabla g = \langle g_x, g_y, g_z\rangle$, or in other words, there exists a scalar λ for which $\nabla f = \lambda \nabla g$. This yields the explicit conditions $f_x = \lambda g_x$, $f_y = \lambda g_y$, $f_z = \lambda g_z$ given above.

- For completeness we also mention that there is an analogous procedure for a problem with two constraints:

- Method (Lagrange Multipliers, 2 constraints): To find the extreme values of $f(x,y,z)$ subject to a pair of constraints $g(x,y,z) = c$ and $h(x,y,z) = d$, define the Lagrange function $L(x,y,z,\lambda,\mu) = f(x,y,z) - \lambda \cdot [g(x,y,z) - c] - \mu \cdot [h(x,y,z) - d]$. Then any extreme value of $f(x,y,z)$ subject to the constraint constraints $g(x,y,z) = c$ and $h(x,y,z) = d$ must occur at a critical point of $L(x,y,z,\lambda,\mu)$.
  - The method also works with more than three variables, and has a natural generalization to more than two constraints. (It is fairly rare to encounter systems with more than two constraints.)

- Example: Find the maximum and minimum values of $f(x,y) = 2x+3y$ subject to the constraint $x^2+4y^2 = 100$.
  - We have $g = x^2 + 4y^2$, and we compute $f_x = 2$, $g_x = 2x$, $f_y = 3$, and $g_y = 8y$.
  - Thus we have the system $2 = 2x\lambda$, $3 = 8y\lambda$, and $x^2 + 4y^2 = 100$.
  - Solving the first two equations gives $x = \dfrac{1}{\lambda}$ and $y = \dfrac{3}{8\lambda}$. Then plugging in to the third equation yields $\left(\dfrac{1}{\lambda}\right)^2 + 4\left(\dfrac{3}{8\lambda}\right)^2 = 100$, so that $\dfrac{1}{\lambda^2} + \dfrac{9}{16\lambda^2} = 100$. Multiplying both sides by $16\lambda^2$ yields $25 = 100(16\lambda^2)$, so that $\lambda^2 = \dfrac{1}{64}$, hence $\lambda = \pm\dfrac{1}{8}$.
  - Thus, we obtain the two points $(x,y) = (8,3)$ and $(-8,-3)$.
  - Since $f(8,3) = 25$ and $f(-8,-3) = -25$, the maximum is $\boxed{f(8,3) = 25}$ and the minimum is $\boxed{f(-8,-3) = -25}$.

- Example: Find the maximum and minimum values of $f(x,y,z) = x + 2y + 2z$ subject to the constraint $x^2 + y^2 + z^2 = 9$.
  - We have $g = x^2 + y^2 + z^2$, and also $f_x = 1$, $g_x = 2x$, $f_y = 2$, $g_y = 2y$, $f_z = 2$, $g_z = 2z$.
  - Thus we have the system $1 = 2x\lambda$, $2 = 2y\lambda$, $2 = 2z\lambda$, and $x^2 + y^2 + z^2 = 9$.

- Solving the first three equations gives $x = \frac{1}{2\lambda}$, $y = \frac{1}{\lambda}$, $z = \frac{1}{\lambda}$; plugging in to the last equation yields $\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 = 9$, so $\frac{9}{4\lambda^2} = 9$, so that $\lambda = \pm\frac{1}{2}$.
- This gives the two points $(x, y, z) = (1, 2, 2)$ and $(-1, -2, -2)$.
- Since $f(1, 2, 2) = 9$ and $f(-1, -2, -2) = -9$, the maximum is $\boxed{f(1, 2, 2) = 9}$ and the minimum is $\boxed{f(-1, -2, -2) = -9}$.

- Example: Maximize the volume $V = \pi r^2 h$ of a cylindrical can given that its surface area $SA = 2\pi r^2 + 2\pi rh$ is $150\pi\,\mathrm{cm}^2$.
  - We have $f(r, h) = \pi r^2 h$ and $g(r, h) = 2\pi r^2 + 2\pi rh$, so $f_r = 2\pi rh$, $g_r = 4\pi r + 2\pi h$, $f_h = \pi r^2$, $g_h = 2\pi r$.
  - Thus we have the system $2\pi rh = (4\pi r + 2\pi h)\lambda$, $\pi r^2 = (2\pi r)\lambda$, and $2\pi r^2 + 2\pi rh = 150\pi$.
  - We clearly cannot have $r = 0$ since that contradicts the third equation, so we can assume $r \neq 0$.
  - Cancelling r from the second equation and then solving for λ yields $\lambda = \frac{r}{2}$. Plugging into the first equation (and cancelling the πs) yields $2rh = (4r + 2h) \cdot \frac{r}{2}$, so dividing by r yields $2h = 2r + h$, so that $h = 2r$.
  - Finally, plugging in $h = 2r$ to the third equation (after cancelling the πs) yields $2r^2 + 4r^2 = 150$, so that $r^2 = 25$ and thus $r = \pm 5$.
  - The two candidate points are $(r, h) = (5, 10)$ and $(-5, -10)$; since we only want positive values we are left only with $(5, 10)$, which by the physical setup of the problem must be the maximum.
  - Therefore, the maximum volume occurs with $r = 5$cm and $h = 10$cm, and is $\boxed{f(5, 10) = 250\pi\mathrm{cm}^3}$.

- Example: An assembly line involving f full-time workers, p part-time workers, and r robots has a total production level of $T(f, p, r) = 80f^{0.7}p^{0.2}r^{0.1}$ gizmos per day. Each full-time worker's compensation totals \$200 per day, each part-time worker's compensation totals \$80 per day, and each robot's maintenance costs total \$40 per day. If the daily operating budget is \$4000, how many of each type of worker, and how many robots, should be employed to maximize total daily production?
  - We wish to maximize the function $T(f, p, r) = 80f^{0.7}p^{0.2}r^{0.1}$ subject to the constraint $200f + 80p + 40r = 4000$, so that $g(f, p, r) = 200f + 80p + 40r$.
  - We may maximize the function T directly, but the resulting calculation is somewhat unpleasant. It is easier to maximize the logarithm of T, namely $\ln(T) = \ln(80) + 0.7\ln(f) + 0.2\ln(p) + 0.1\ln(r)$, instead.
  - Taking the partial derivatives then yields the system $\frac{0.7}{f} = 200\lambda$, $\frac{0.2}{p} = 80\lambda$, $\frac{0.1}{r} = 40\lambda$, and $200f + 80p + 40r = 4000$.
  - The first three equations yield $f = \frac{7}{2000\lambda}$, $p = \frac{1}{400\lambda}$, and $r = \frac{1}{400\lambda}$.
  - Plugging these expressions into the last equation then yields $200 \cdot \frac{7}{2000\lambda} + 80 \cdot \frac{1}{400\lambda} + 40 \cdot \frac{1}{400\lambda} = 4000$, which simplifies to $\frac{1}{\lambda} = 4000$ and thus $\lambda = \frac{1}{4000}$. This yields a unique candidate triple $(f, p, r) = (14, 10, 10)$, which by the setup of the problem must be a maximum.
  - We conclude that the maximum production occurs with $\boxed{\text{14 full-time workers, 10 part-time workers, and 10 robots}}$, and is $T(14, 10, 10) \approx 1012.46$ gizmos per day.
  - Remark: If we tried to maximize T directly, the system of equations would be $56f^{-0.3}p^{0.2}r^{0.1} = 200\lambda$, $16f^{0.7}p^{-0.8}r^{0.1} = 80\lambda$, $8f^{0.7}p^{0.2}r^{-0.9} = 40\lambda$, $200f + 80p + 40r = 4000$. This system is not nearly as easy to solve as the one above; one approach is to divide the second and third equations by the first one, then solve for two of f, p, r in terms of the other one, and finally plug in to the last equation.

Well, you're at the end of my handout. Hope it was helpful.