Minimum and maximum values, Lagrange multipliers

In a process very similar to that in single-variable calculus, we look for minima and maxima by finding points where the first derivatives are zero:

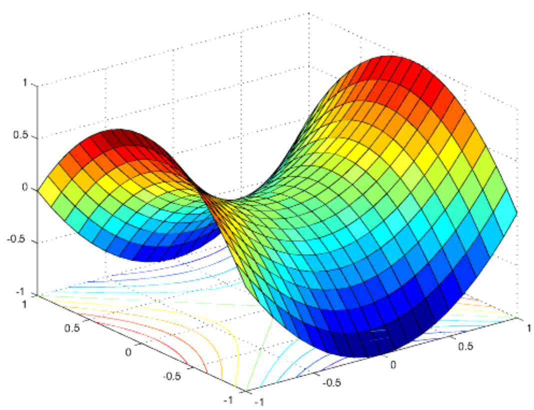
$f_x = \frac{\partial f}{\partial x} = 0$ and $f_y = \frac{\partial f}{\partial y} = 0$

These are points where the tangent plane is parallel to the xy plane.

A *critical point* is defined as a point where $f_x = 0$ and $f_y = 0$, or where one of these does not exist. However, a critical point need not be a maximum or minimum. Take, for example, the curve:
$z = y^2 - x^2$

$f_x = -2x$ and $f_y = 2y$

so the one critical point is at (0,0)



However, as seen in the graph above, this point is neither a minimum nor maximum, but rather a *saddle point*.

To test critical points for being a maximum, minimum, or neither, we use something similar to the second derivative test in single-variable calculus.

First we define the function:

$D(a,b) = f_{xx}(a,b)\cdot f_{yy}(a,b) - [f_{xy}(a,b)]^2$ such that

1. If $D(a,b) > 0$ and $f_{xx}(a,b) > 0$, then the point is a minimum
2. If $D(a,b) > 0$ and $f_{xx}(a,b) < 0$, then the point is a maximum
3. If $D(a,b) < 0$, the point is neither and is instead a saddle point
4. If $D(a,b) = 0$, the nature of the point is unresolved

Where this set of rules comes-from is not particularly instructive, so I will leave it for an appendix at the end of the notes.

As an example, let's find the critical points for $f(x,y) = x^2 + xy + y^2 + y$ and determine if they are maxima, minima, or saddle points.

First, find the partial derivatives and set them equal to zero:
$f_x = 2x + y = 0$
$f_y = x + 2y + 1= 0$
Has only one solution at $(\frac{1}{3}, \frac{-2}{3})$.

Then use the second derivatives test:

$f_x = 2x + y$ and $f_{xx} = 2$ and $f_{xy} = 0$
$f_y = x + 2y$ and $f_{yy} = 2$

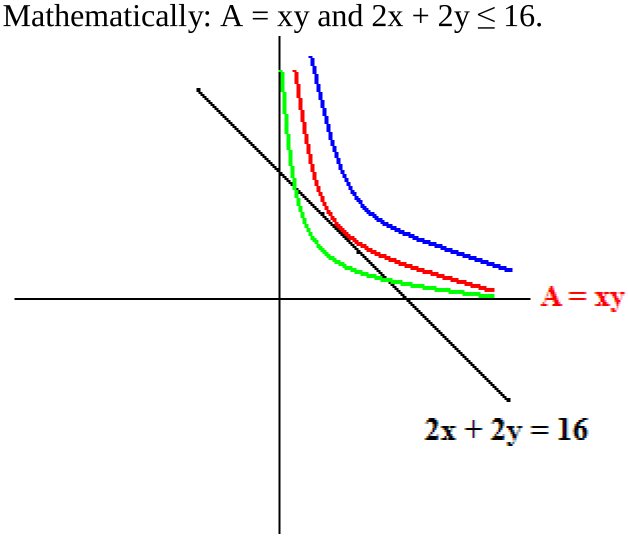
$D(a,b) = f_{xx}(a,b) \cdot f_{yy}(a,b) - [f_{xy}(a,b)]^2 = 4 – 0 = 4$
$f_{xx} = 2$

$D > 0$ and $f_{xx} > 0$, therefore the point $(\frac{1}{3}, \frac{-2}{3})$ is a minimum at a value of $-\frac{1}{3}$.

To find the absolute maxima and minima within a given range of values, simply find the values at all critical points within the range and all values along the boundaries of the range.

Now suppose we want to maximize some function, $f(x,y,z)$ with the restraining condition that $g(x,y,z) = k$. For instance, suppose we wanted to maximize the volume of a six-sided box with sides x, y, and z following the restraint that the total surface area used for the sides is not greater than k. Mathematically: $V = xyz$ and $(2xy + 2yz + 2xz) \leq k$.

To help visualize a geometric solution, imagine the simpler case of two variables: We want to maximize the area of a piece of paper with the constraint that the perimeter is not greater than 16. Mathematically: $A = xy$ and $2x + 2y \leq 16$.



That means there is going to be some curve (out of a series of possible curves, some higher values of A, some lower) that just touches the line of $2x + 2y = 16$ and satisfies that condition while being the highest value curve of A. At that point of contact between the two curves, they share a common slope and thus also a common angle perpendicular to the slope (also known as the normal to the curve). In the previous notes, we saw that the gradient vector is perpendicular to the tangent vector, so the two curves have a gradient vector pointed in the same direction (although they may differ by a scalar factor). Mathematically, this would be:

$$\nabla A(x,y) = \lambda \cdot \nabla P(x,y)$$

where A is area, P is perimeter, and λ is that scalar factor called a *Lagrange multiplier*.

Extending these ideas into three dimensions (and using general functions f and g), we have:

$$\nabla f(x,y,z) = \lambda \cdot \nabla g(x,y,z)$$

Here the visual would be a curved surface that follows the function f(x,y,z) and, in the same space, a series of level planes that follow the function g(x,y,z). The maximum value of f that satisfies the constraint g will occur at the point where the surface of f just touches one of the planes of g.

As an example, let’s return to the example of the six-sided box and give the restriction that the total surface area cannot be greater than 12. We then have:

$$V = xyz \qquad \text{and} \qquad A = 2xy + 2yz + 2xz$$

Using the equation with the Lagrange multiplier, we have:

$$\nabla f(x,y,z) = \lambda \cdot \nabla g(x,y,z) \qquad \text{or} \qquad \nabla V(x,y,z) = \lambda \cdot \nabla A(x,y,z)$$

which can be broken into components:

$$V_x = \lambda \cdot A_x \qquad \text{or} \qquad yz = \lambda \cdot (2y + 2z)$$
$$V_y = \lambda \cdot A_y \qquad \text{or} \qquad xz = \lambda \cdot (2x + 2z)$$
$$V_z = \lambda \cdot A_z \qquad \text{or} \qquad xy = \lambda \cdot (2y + 2x)$$

with the further condition that: $12 = 2xy + 2yz + 2xz$

Altogether, we have four equations with four unknowns, which is a solvable system of equations. Such algebra is often very complicated, but here it is simple. Because of the symmetry of the first three equations, we can know that $x = y = z$ and the fourth equations becomes:

$$12 = 2x^2 + 2x^2 + 2x^2 \qquad \text{and} \qquad x = y = z = \sqrt{2}$$

where λ can be left unknown, though it here happens to be $\frac{1}{2\sqrt{2}}$

Lastly, we can make one further extension and suppose that the function f(x,y,z) has two constraints of g(x,y,z) = k and h(x,y,z) = c. Then we simply use two Lagrange multipliers and solve for five unknowns:

$\nabla f(x,y,z) = \lambda \cdot \nabla g(x,y,z) + \mu \cdot \nabla h(x,y,z)$

Geometrically, the two constraints would form two level surfaces that intersect at some curve, *C*. The maximum value of f will occur along this curve at point *P* and satisfy the conditions that ∇f is perpendicular to *C* and so are ∇g and ∇h.

Appendix:

If we take the second order directional derivative of f in the direction of the unit vector, **u** = <h,k>, we have:

$D_u f = f_x h + f_y k$

$D_u^2 f = D_u(D_u)f = D_u(f_x h + f_y k) = (f_{xx}h + f_{yx}k)h + (f_{xy}h + f_{yy}k)k = f_{xx}h^2 + 2f_{xy}hk + f_{yy}k^2$

Completing the square, this becomes: $D_u^2 f = f_{xx}(h + \frac{f_{xy}}{f_{xx}}k)^2 + \frac{k^2}{f_{xx}}(f_{xx}f_{yy} - f_{xy}^2)$

Condition 1 states that, "If D(a,b) > 0 and $f_{xx}(a,b) > 0$, then the point is a minimum". This is because both the first and second terms of $D_u^2 f$ are positive, making the second derivative positive and the curvature upward.

Condition 2 states that, "If D(a,b) > 0 and $f_{xx}(a,b) < 0$, then the point is a maximum". This is because both the first and second terms of $D_u^2 f$ are negative, making the second derivative negative and the curvature downward.

Condition 3 states that, "If D(a,b) < 0, the point is neither and is instead a saddle point". This is because now the first and second terms have opposite signs and will curve up or down depending upon the value of f_{xx}.