Here are some notes on these two topics.

**The second derivative test**

Suppose we are given a function \( f(x, y) \) of two variables and we wish to find the local max/min points. We describe the test to be used and why it works, when it does.

1. First recall that the directional derivative of \( f \) in a direction \( u \) and at the point \( x, y \) is given by the formula
   \[
   D_u(f) = \nabla f(x, y) \cdot u/|u|.
   \]
   Also, if this is not positive for some direction \( u \), then the function values get bigger in the \( u \)-direction and get smaller in the \( -u \)-direction. Thus the function could not have a local max/min at the point \((x, y)\). Thus, we have a necessary condition for local max/min, namely the point should be critical, i.e. \( f_x = 0 = f_y \). (Actually, we should say that either these are zero or are undefined, but we assume for convenience that we can find all the necessary derivatives, in other words, \( f(x, y) \) is sufficiently differentiable.)

2. Now suppose we have a critical point and we wish to have a local minimum. This means that the function should increase in all directions. To check this, fix a direction \( u = (h, k) \) and study the function \( f(x, y) \) restricted to the line \(<x + ht, y + kt>\). Thus define:
   \[
   F(t) = f(x + ht, y + kt).
   \]
   What we are asking is that this function have a local minimum at \( t = 0 \) regardless of our choice of \( h, k \). Our old second derivative test applied to this function of one variable \( t \) gives that
   \[
   F'(0) = 0 \quad \text{and} \quad F''(0) > 0.
   \]
   If we calculate these, we get
   \[
   F'(t) = f_x(x + ht, y + kt)h + f_y(x + ht, y + kt)k \quad \text{so} \quad F'(0) = f_x(x, y)h + f_y(x, y)k
   \]
   Or briefly,
   \[
   F'(0) = hf_x + kf_y.
   \]
   Thus \( F'(0) = 0 \) for all \( h, k \) requires that \( f_x = f_y = 0 \). Now
   \[
   F''(0) = h(f_{xx}h + f_{xy}k) + k(f_{yx}h + f_{yy}k) = f_{xx}h^2 + 2f_{xy}hk + f_{yy}k^2.
   \]
   We want to see when this can be claimed to be positive for all \((h, k)\). First, let \( k = 0 \) and we see that we need \( f_{xx} > 0 \). Now, we consider \( k \neq 0 \) and it is easy to see that we can factor out a \( k^2 \) to write
   \[
   F''(0) = k^2(f_{xx}w^2 + 2f_{xy}w + f_{yy}).
   \]
   where we have written \( w \) for the ratio \( h/k \). Thus we want to see when the quadratic function \( f_{xx}w^2 + 2f_{xy}w + f_{yy} \) comes out positive for all values of \( w \). An obvious criterion is that the quadratic equation
   \[
   f_{xx}w^2 + 2f_{xy}w + f_{yy} = 0
   \]
   has no roots and the good old quadratic formula says that we must have the discriminant \( 4f_{xy}^2 - 4f_{xx}f_{yy} \) is negative. This gives us \( D = f_{xx}f_{yy} - f_{xy}^2 > 0 \) as a condition, after some algebraic manipulation! Indeed, if the quantity \( D < 0 \) then the quadratic equation will have distinct real roots and we will indeed get different signs for our \( F''(0) \) depending on the choice of \( h, k \). In case, \( D = 0 \), we will get that the second derivative is always bigger than or equal to zero. This is like a local minimum, except in a certain direction, the function stays constant. To summarize:
If \( D > 0 \) and \( f_{xx} > 0 \), then we have a local minimum. If \( D = 0 \) and \( f_{xx} < 0 \), then we have a local minimum in all directions, except one and along this, the function is constant. If \( D < 0 \), then we don’t have a local minimum (or maximum). Clearly a similar argument can be made to deduce that we have a local maximum if \( D > 0 \) and \( f_{xx} < 0 \) and a conditional local maximum if \( D = 0 \) and \( f_{xx} < 0 \), which means that we have a local maximal along all but one directions and along that direction we have a constant function. Also, we can repeat the above tests with the roles of \( x, y \) interchanged. What this means is that in case of \( D = 0 \), all we need is that one of the two derivatives \( f_{xx}, f_{yy} \) be nonzero. This is the complete second derivative test. Note that if \( D = 0 \) and all the second partials are zero, then we have no conclusion from this test!

3. The complete test. Let \( D = f_{xx}f_{yy} - f_{xy}^2 \).

- If the point \( x, y \) is critical and \( D < 0 \) then it is a saddle point, meaning the directional derivatives can be both positive and negative if we choose different directions.
- If \( D \geq 0 \) and \( f_{xx} \) or \( f_{yy} \) is positive, then it is a local minimum. Moreover, in case \( D = 0 \), this is a conditional minimum.
- If \( D \geq 0 \) and \( f_{xx} \) or \( f_{yy} \) is negative, then it is a local maximum. Moreover, in case \( D = 0 \), this is a conditional maximum.

4. Using the function \( F(t) \) we can develop suitable higher order approximations for the function as follows. The linear approximation to \( F(t) \) at \( t = 0 \) gives \( L(t) = F(0) + F'(0)t = f(x, y) + (hf_x + kf_y)t \). If we set \( \Delta x = ht \) and \( \Delta y = kt \), we get the usual linear approximation for \( f \), namely, \( f(x, y) + f_x \Delta x + f_y \Delta y \). If we use the second order approximation for \( F(t) \) at \( t = 0 \), then we have \( Q(t) = F(0) + F'(0)t + (1/2)F''(0)t^2 \) which will reduce to the expression:

\[
f(x, y) + f_x \Delta x + f_y \Delta y + (1/2)(f_{xx} \Delta x^2 + 2f_{xy} \Delta x \Delta y + f_{yy} \Delta y^2).
\]

We can also develop higher order approximations as well as suitable Taylor expansions of various orders. Indeed, nothing limits us to the case of two variables. Similar formulae can be made and are valid for more variables, as long as the function has enough derivatives.

**The Lagrange Multipliers**

1. In the above analysis, we learned how to find local extrema of a function by finding critical points, or points at which all directional derivatives are zero. In many cases, the second derivative test can be omitted, if we can argue the extreme nature of the function at the point indirectly. So we wish to learn efficient methods for finding critical points.

However, in real life situations, we need to find the critical points when the function is restricted to a surface or a curve. What this means is that the available directional derivatives are restricted - since we have to live on a surface or a curve. The Lagrange Multipliers give a very efficient method for finding such critical points. Usually, there is no equivalent second derivative test though.

2. A typical situation is when we wish to find critical points for a function \( f \) subject to a restriction \( g = 0 \). Then the available directions are all directions in the tangent plane to \( g = 0 \). Since \( \nabla g \) is a normal to the tangent plane, we see that the available directions \( u \) satisfy \( \nabla g \cdot u = 0 \). Thus we want to impose the condition that

\[
\text{Whenever } \nabla g \cdot u = 0 \text{ we have } \nabla f \cdot u = 0.
\]
It is clear that this will be true if $\nabla f$ is a scalar multiple of $\nabla g$ and it is not hard to convince yourself that this condition is necessary as well! So, we get that a critical point occurs when

$$\nabla f - \lambda \nabla g = 0 \quad \text{and} \quad g = 0.$$ 

3. Here is an example. Let us find the distance between the point $(a, b, c)$ and a plane $px + qy + rz - s = 0$. We take the square of the distance between $(a, b, c)$ and a point $(x, y, z)$ on the plane as our function. In other words we have to minimize $f = (x - a)^2 + (y - b)^2 + (z - c)^2$ subject to the condition $px + qy + rz - s = 0$.

So we have to solve the equations: $<2(x - a), 2(y - b), 2(z - c) > - \lambda < p, q, r > = 0$ and $px + qy + rz = 0$.

These are actually four equations

$$x - a = \lambda p/2, \quad y - b = \lambda q/2, \quad z - c = \lambda r/2 \quad \text{and} \quad px + qy + rz = s.$$ 

Solving and the first three, we get a single critical point and plugging it into the fourth, we evaluate the unknown $\lambda$. You should check that this yields:

$$x = a + \lambda p/2, \quad y = b + \lambda q/2, \quad z = c + \lambda r/2, \quad p(a + \lambda p/2) + q(b + \lambda q/2) + r(c + \lambda r/2) = s.$$ 

This gives

$$\lambda = \frac{-2(pa + qb + rc)}{p^2 + q^2 + r^2}.$$ 

Moreover, if we substitute in $f$, we get

$$f = \lambda^2(p^2 + q^2 + r^2)/4 = \frac{(pa + qb + rc - s)^2}{p^2 + q^2 + r^2}.$$ 

This gives the formula for the actual distance (square root of $f$)

$$\frac{|pa + qb + rc - s|}{\sqrt{p^2 + q^2 + r^2}}.$$ 

4. How do we know that this is the actual absolute minimum distance? From geometry, we know that there is a minimum and there is no maximum since the distances can obviously get arbitrarily large. Since we found only one critical point, it must give us the absolute minimum!

The final formula is quite pleasant. We simply substitute the coordinates of the point in the equation of the plane and divide by the length of the normal vector $<p, q, r>$. In other words, just the substitution will give the distance, if the equation of the plane is fixed up so its normal vector is a unit vector.

Indeed, if we skip the absolute value sign, then we get the signed distance which also informs us about which side of the plane does the point live on. All points on one side of the plane always give the same sign! One way of fixing this is to make $s < 0$ by multiplying the equation by $-1$ if needed. Then the sign is positive if and only if the point is on the same side as the origin!

This formula is well worth learning for its own sake. It works beautifully in higher dimensions as well and indeed in the plane too - though you probably did not study it! The Lagrange Multiplier calculations work the same way in all dimensions.
5. Many of the problems in this section are best handled by clever substitutions rather than a brute force calculation. More on this when I illustrate more problems in class. The more problems you solve with proper thought, the better you will get it this game!

Place for your notes and reminders