MATH 20C Lecture 15 - Monday, February 10, 2014

Recall that $D_{\hat{u}}f = r\nabla f \cdot \hat{u} = |\nabla f| cos\theta$. Maximal for $\cos\theta = 1$, when $\hat{u}$ is in direction of ∇f. Hence: direction of ∇f is that of fastest increase of f, and $|\nabla f|$ is the directional derivative in that direction.
It's minimal in the opposite direction.
We have $D_{\hat{u}}f = 0$ when $\hat{u} \perp \nabla f$, i.e. when $\hat{u}$ is tangent to direction of level surface.

Example Compute the directional derivative of $f = xe^{-yz}$ at $P = (1, 2, 0)$ in the direction of $\vec{v} = \langle 1, 1, 1\rangle$.

$\nabla f = \langle e^{-yz}, -xze^{-yz}, -xye^{-yz}\rangle$ so $\nabla f(P) = \langle 1, 0, -2\rangle$. The unit vector in the direction of $\vec{v}$ is $\hat{u} = \vec{v}/|\vec{v}| = \langle 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}\rangle$. So $D_{\hat{u}}f(P) = \nabla f(P) \cdot \hat{u} = -1/\sqrt{3}$. Therefore f is decreasing in the direction of $\vec{v}$.
In which direction is f increasing fastest? ($\nabla f(P) = \langle 1, 0, -2\rangle$) Decreasing fastest? $-\nabla f(P) = \langle -1, 0, 2\rangle$) Tangent plane at $(1, 2, 0)$?($x - 2z = 1$)

Implicit differentiation

Example: $x^2 + yz + z^3 = 8$. Viewing $z = z(x, y)$, compute $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$.

Take $\frac{\partial}{\partial x}$ of both sides of $x^2 + yz + z^3 = 8$. Get $2x + y\frac{\partial z}{\partial x} + 3z^2\frac{\partial z}{\partial x} = 0$, hence $\frac{\partial z}{\partial x} = -\frac{2x}{y+3z^2} = -\frac{2}{3}$.

In general, consider a surface $F(x, y, z) = c$. The we can view $z = z(x, y)$ as a function of two independent variables x, y and compute $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$. To do so, we take the partial derivative with respect to x of both sides of the equation $F(x, y, z) = c$ and get (by the chain rule)

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0.$$

But $\partial x/\partial x = 1$ and, since x and y are independent, $\partial y/\partial x = 0$ (changing x does not affect y). Hence the equation above really says that $F_x + F_z\frac{\partial z}{\partial x} = 0$ which implies

$$\boxed{\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}.}$$

Similarly,

$$\boxed{\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}.}$$

Changing gears, let's see how we can recover f from its gradient. Say $\nabla f = \langle 3x^2y, x^3 + 2z, 2y + \cos z\rangle$. We proceed by successive integration. We are given that $f_x = 3x^2y$. Integrating with respect to x (view y, z as constants), we see that $f = x^3y + g(y, z)$. Therefore

$$f_y = x^3 + \frac{\partial g}{\partial y}.$$

But we know from the gradient that $f_y = x^3 + 2z$, hence $g_y = 2z$. Integrate with respect to y and get $g = 2yz + h(z)$, hence $f = x^3y + 2yz + h(z)$. Since $f_z = 2y + \cos z$ we get that $\frac{dh}{dz} = \cos z$, so $h(z) = \sin z + C$. Substituting in the expression of f gives $f = x^3y + 2yz + \sin z + C$.

Min/max in several variables

At a local max or min, $f_x = 0$ and $f_y = 0$ (since (x_0, y_0) is a local max or min of the slice). Because 2 lines determine tangent plane, this is enough to ensure that the tangent plane is horizontal.
Definition A critical point of f is a point (x_0, y_0) where $f_x = 0$ and $f_y = 0$. A critical point may be a local min, local max, or saddle. Or degenerate. Pictures shown of each type.

MATH 20C Lecture 16 - Wednesday, February 12, 2014

Min/max in several variables

Recall that a critical point of f is a point (x_0, y_0) where $\nabla f(x_0, y_0) = 0$.

A critical point may be a local min, local max, or saddle. Or degenerate. To decide, apply **second derivative test**.
Example: $f(x, y) = x^2 - 2xy + 3y^2 + 2x - 2y$. Critical point: $f_x = 2x - 2y + 2 = 0$, $f_y = -2x + 6y - 2 = 0$, gives $(x_0, y_0) = (-1, 0)$ (only one critical point).
Definition The hessian matrix of f is

$$H(x, y) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial y \partial x} \\ \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

Second derivative test

Let (x_0, y_0) be a critical point of f.

Case 1 $\det H > 0, f_{xx} > 0$: (x_0, y_0) is a local minimum

Case 2 $\det H > 0, f_{xx} < 0$: (x_0, y_0) is a local maximum

Case 3 $\det H < 0$: (x_0, y_0) is a saddle point

Case 4 $\det H = 0$: cannot tell (need higher order derivatives)

Example 1 Find the local min/max of $f(x, y) = x + y + \frac{1}{xy}$.

Step 1 *Find critical points by solving the* 2×2 *system of equations*

$$\begin{cases} f_x = 0 \\ f_y = 0 \end{cases}$$

In this case, the system is

$$\begin{cases} \frac{1}{x^2 y} = 1 \\ \frac{1}{xy^2} = 1. \end{cases}$$

Divide the first equation by the second and get $x = y$, plug back into the first equation and get $x^3 = 1$. So the only critical point is $(1, 1)$.

Showed slide asking students if this point is a local max/min or saddle. Most got it right (local min). Now let's do it rigorously.

Step 2 *Compute the Hessian matrix*

$$H(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}.$$

Recall that $f_{xy} = f_{yx}$.

In our case, get $H(x,y) = \begin{bmatrix} \frac{2}{x^3 y} & \frac{1}{x^2 y^2} \\ \frac{1}{x^2 y^2} & \frac{2}{xy^3} \end{bmatrix}$.

Step 2 *Compute the Hessian matrix at each of the critical points.*

$$H(1,1) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Step 4 *Apply the second derivative test for each critical point.*

$\det H(1,1) = 4 - 1 = 3 > 0$ and $f_{xx} = 2 > 0$, so $(1, 1)$ is a local minimum.

Attention! We can also infer the nature of a critical point from the contour plot. Showed picture and discussed possibilities. Most students got the right answer.

Example 2 $f(x,y) = (x^2 + y^2)e^{-x}$

Step 1 *Find critical points by solving the* 2×2 *system of equations*

$$\begin{cases} f_x = 0 \\ f_y = 0 \end{cases}$$

In this case, the system is

$$\begin{cases} (2x - x^2 - y^2)e^{-x} = 0 \\ 2ye^{-x} = 0. \end{cases}$$

The second equation tells us that $y = 0$. Plug back into the first equation and get $x^2 - 2x = 0$. So critical points are $(0, 0)$ and $(2, 0)$.

Step 2 *Compute the Hessian matrix*

$$H(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}.$$

In our case, get $H(x,y) = \begin{bmatrix} (2 - 4x + x^2 + y^2)e^{-x} & -2ye^{-x} \\ -2ye^{-x} & 2e^{-x} \end{bmatrix}$.

Step 2 *Compute the Hessian matrix at each of the critical points.*

$$H(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

and

$$H(2,0) = \begin{bmatrix} -2e^{-2} & 0 \\ 0 & 2e^{-2} \end{bmatrix}.$$

Step 4 *Apply the second derivative test for each critical point.*

- For $(0,0)$: $\det H(0,0) = 4 > 0$ and $f_{xx} = 2 > 0$, so $(0,0)$ is a local minimum.
- For $(2,0)$: $\det H(2,0) = -4e^{-4} < 0$, so $(2,0)$ is a saddle point.

NOTE: the global min/max of a function is not necessarily at a critical point! Need to check boundary / infinity.

In Example 2 above, to find the global min/max of f in the region $0 \leq x, y \leq 1$, w need to check what happens on the boundary. Namely we have to look at $f(0,y), f(1,y), f(x,0)$ and $f(x,1)$. We have to compute the min/max for these 4 functions and compare to the value at critical points inside the square (if any).

MATH 20C Lecture 17 - Friday, February 14, 2014

Lagrange multipliers

Problem: min/max of a function $f(x,y,z)$ when variables are constrained by an equation $g(x,y,z) = c$.

Example: find point of the curve $xy^2 = 4$ closest to origin. I.e. minimize $\sqrt{x^2+y^2}$, or better $f(x,y) = x^2 + y^2$, subject to $g(x,y) = xy^2 = 4$. Shown picture.

Observe on picture: at the minimum, the level curves are tangent to each other, so the normal vectors ∇f and ∇g are parallel.

So: there exists λ ("multiplier") such that $\nabla f = \lambda \nabla g$.

We replace the constrained min/max problem in 2 variables with 3 equations involving 3 variables x, y, λ :

$$\begin{cases} f_x = \lambda g_x \\ f_y = \lambda g_y \\ g(x,y) = c \end{cases} \qquad \text{i.e. in our case} \qquad \begin{cases} 2x = \lambda y^2 \\ 2y = 2\lambda xy \\ xy^2 = 4. \end{cases}$$

Since we need $x, y \neq 0$ in the third equation, the second equation yields $x = 1/\lambda$. Plugging this into the first equation tells us that $y = \pm\sqrt{2}/\lambda$. Substituting x and y in the third equation, we get $\lambda^3 = 1/2$. Get $\lambda = 1/\sqrt[3]{2}, x = \sqrt[3]{2}, y = \pm\sqrt[6]{32}$. Get two points where the distance to origin is minimal, $(\sqrt[3]{2}, \sqrt[6]{32})$ and $(\sqrt[3]{2}, -\sqrt[6]{32})$.

Warning: method doesn't say whether we have a min or a max, and second derivative test DOES NOT apply with constrained variables. Need to answer using geometric argument or by comparing values of f.

Why does this work

We are looking for points on $g = c$ where f attains a local min/max when restricted to the level set $g = c$. This means that we want the directional partial derivative $D_{\hat{\mathbf{u}}} f = 0$ for all $\hat{\mathbf{u}}$ unit vectors tangent to $g = c$. Since $0 = D_{\hat{\mathbf{u}}} f = \nabla f \cdot \hat{\mathbf{u}}$ this means that we want ∇f to be perpendicular to the level set $g = c$. We already have a vector tangent to the level set, namely ∇g. So we want $\nabla f \parallel \nabla g$.

Example: Find the min/max of $f(x, y, z) = 3x + y + 4z$ on the surface $x^2 + 3y^2 + 6z^2 = 1$.

Step 1 *Compute the two gradients ∇f and ∇g.*

$\nabla f = \langle 3, 1, 4\rangle \qquad \nabla g = \langle 2x, 6y, 12z\rangle$

Step 2 *Write down the Lagrange multiplier equations $\nabla f = \lambda \nabla g$ and the constraint $g = c$.*

$$\nabla f = \lambda \nabla g \quad \Longrightarrow \quad \begin{cases} 3 &= 2\lambda x \\ 1 &= 6\lambda y \\ 4 &= 12\lambda z \end{cases}$$

$$g = c \quad \Longrightarrow \quad x^2 + 3y^2 + 6z^2 = 1$$

Step 3 *Solve the system, i.e. find points (x, y, z) that satisfy the equations from Step 2.*

WARNING! There is no general method to solve these equations. In each case, you have to think about them and come up with a method. Sometimes it will be impossible to solve without using a computer. (Not on the exam though!)

In our example, note that λ cannot be 0. From the first three equations get $x = \frac{3}{2\lambda}, y = \frac{1}{6\lambda}, z = \frac{1}{3\lambda}$. Substitute these values in the constraint equation and get $\lambda^2 = 3$, so $\lambda = \pm\sqrt{3}$. Therefore there are two points $\left(\frac{\sqrt{3}}{2}, \frac{1}{6\sqrt{3}}, \frac{1}{3\sqrt{3}}\right)$ and $\left(-\frac{\sqrt{3}}{2}, -\frac{1}{6\sqrt{3}}, -\frac{1}{3\sqrt{3}}\right)$ on the given surface at which the gradients of f and g are parallel.

Step 4 *Plug the points you found into f and compare values.*

WARNING! There is no general method to tell you if the points you got are local min/max/saddle. The only way is to plug in points *on the surface/curve* $g = c$ and compare the values of f.

However, sometimes we can find some geometric reason for which the function f would be forced to have a min or a max on the surface/curve $g = c$. For instance, if $g = c$ is closed, then f will have both min and max along $g = c$.

This happens to be the case for us, since $x^2 + 3y^2 + 6z^2 = 1$ describes the shell of an ovoid in 3-space.

$$f\left(\frac{\sqrt{3}}{2}, \frac{1}{6\sqrt{3}}, \frac{1}{3\sqrt{3}}\right) = 3\frac{\sqrt{3}}{2} + \frac{1}{6\sqrt{3}} + 4\frac{1}{3\sqrt{3}} = \frac{2}{\sqrt{3}}.$$

$$f\left(-\frac{\sqrt{3}}{2}, -\frac{1}{6\sqrt{3}}, -\frac{1}{3\sqrt{3}}\right) = -3\frac{\sqrt{3}}{2} - \frac{1}{6\sqrt{3}} - 4\frac{1}{3\sqrt{3}} = -\frac{2}{\sqrt{3}}.$$

The first point is a max, the second is a min.

Application of min/max problems: least squares method

Set up problem: given experimental data (x_i, y_i) $(i = 1, \ldots, n)$, want to find a best-fit line $y = ax+b$ (the unknowns here are a, b, not x, y!) Deviations: $y_i - (ax_i + b)$; want to minimize the total square deviation $D(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$.

$\frac{\partial D}{\partial a} = 0$ and $\frac{\partial D}{\partial b} = 0$ leads to a 2×2 linear system for a and b

$$\begin{aligned}
\left(\sum_{i=1}^{n} x_i^2\right) a + \left(\sum_{i=1}^{n} x_i\right) b &= \sum_{i=1}^{n} x_i y_i \\
\left(\sum_{i=1}^{n} x_i\right) a + nb &= \sum_{i=1}^{n} y_i
\end{aligned}$$

The least-squares setup also works in other cases: e.g. exponential laws $y = ce^{ax}$ (taking logarithms: $\ln y = ax + \ln c$, so setting $b = \ln c$ we reduce to linear case); or quadratic laws $y = ax^2 + bx + c$ (minimizing total square deviation leads to a 3×3 linear system for a, b, c). Example: Moores Law (number of transistors on a computer chip increases exponentially with time): showed picture.