18.02 Multivariable Calculus
Fall 2007

18.02 Lecture 11. – **Tue, Oct 2, 2007**

Differentials.

Recall in single variable calculus: $y = f(x) \Rightarrow dy = f'(x)\,dx$. Example: $y = \sin^{-1}(x) \Rightarrow x = \sin y$, $dx = \cos y\,dy$, so $dy/dx = 1/\cos y = 1/\sqrt{1-x^2}$.

Total differential: $f = f(x, y, z) \Rightarrow df = f_x\,dx + f_y\,dy + f_z\,dz$.

This is a new type of object, with its own rules for manipulating it (df is not the same as Δf! The textbook has it wrong.) It encodes how variations of f are related to variations of x, y, z. We can use it in two ways:

1. as a placeholder for approximation formulas: $\Delta f \approx f_x \Delta x + f_y \Delta y + f_z \Delta z$.

2. divide by dt to get the **chain rule**: if $x = x(t)$, $y = y(t)$, $z = z(t)$, then f becomes a function of t and $\dfrac{df}{dt} = f_x \dfrac{dx}{dt} + f_y \dfrac{dy}{dt} + f_z \dfrac{dz}{dt}$

Example: $w = x^2 y + z$, $dw = 2xy\,dx + x^2\,dy + dz$. If $x = t$, $y = e^t$, $z = \sin t$ then the chain rule gives $dw/dt = (2te^t)\,1 + (t^2)\,e^t + \cos t$, same as what we obtain by substitution into formula for w and one-variable differentiation.

Can justify the chain rule in 2 ways:

1. $dx = x'(t)\,dt$, $dy = y'(t)\,dt$, $dz = z'(t)\,dt$, so substituting we get $dw = f_x dx + f_y dy + f_z dz = f_x\,x'(t)\,dt + f_y\,y'(t)\,dt + f_z\,z'(t)\,dt$, hence dw/dt.

2. (more rigorous): $\Delta w \simeq f_x \Delta x + f_y \Delta y + f_z \Delta z$, divide both sides by Δt and take limit as $\Delta t \to 0$.

Applications of chain rule:

Product and quotient formulas for derivatives: $f = uv$, $u = u(t)$, $v = v(t)$, then $d(uv)/dt = f_u\,u' + f_v\,v' = vu' + uv'$. Similarly with $g = u/v$, $d(u/v)/dt = g_u\,u' + g_v\,v' = (1/v)\,u' + (-u/v^2)\,v' = (u'v - uv')/v^2$.

Chain rule with more variables: for example $w = f(x, y)$, $x = x(u, v)$, $y = y(u, v)$. Then $dw = f_x\,dx + f_y\,dy = f_x(x_u du + x_v dv) + f_y(y_u du + y_v dv) = (f_x x_u + f_y y_u)\,du + (f_x x_v + f_y y_v)\,dv$. Identifying coefficients of du and dv we get $\partial f/\partial u = f_x x_u + f_y y_u$ and similarly for $\partial f/\partial v$. It's not legal to "simplify by ∂x".

Example: polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$. Then $f_r = f_x x_r + f_y y_r = \cos\theta\,f_x + \sin\theta\,f_y$, and similarly f_θ.

18.02 Lecture 12. – **Thu, Oct 4, 2007**

Handouts: PS4 solutions, PS5.

Gradient.

Recall chain rule: $\dfrac{dw}{dt} = w_x \dfrac{dx}{dt} + w_y \dfrac{dy}{dt} + w_z \dfrac{dz}{dt}$. In vector notation: $\dfrac{dw}{dt} = \nabla w \cdot \dfrac{d\vec{r}}{dt}$.

Definition: $\nabla w = \langle w_x, w_y, w_z \rangle$ – GRADIENT VECTOR.

Theorem: ∇w is perpendicular to the level surfaces $w = c$.

Example 1: $w = ax + by + cz$, then $w = d$ is a plane with normal vector $\nabla w = \langle a, b, c \rangle$.

Example 2: $w = x^2 + y^2$, then $w = c$ are circles, $\nabla w = \langle 2x, 2y \rangle$ points radially out so $\perp$ circles.

Example 3: $w = x^2 - y^2$, shown on applet (Lagrange multipliers applet with g disabled).

∇w is a vector whose value depends on the point (x,y) where we evaluate w.

Proof: take a curve $\vec r = \vec r(t)$ contained inside level surface $w = c$. Then velocity $\vec v = d\vec r/dt$ is in the tangent plane, and by chain rule, $dw/dt = \nabla w \cdot d\vec r/dt = 0$, so $\vec v \perp \nabla w$. This is true for every $\vec v$ in the tangent plane.

Application: tangent plane to a surface. Example: tangent plane to $x^2+y^2-z^2=4$ at $(2,1,1)$: gradient is $\langle 2x, 2y, -2z\rangle = \langle 4,2,-2\rangle$; tangent plane is $4x+2y-2z=8$. (Here we could also solve for $z=\sqrt{x^2+y^2-4}$ and use linear approximation formula, but in general we can't.)

(Another way to get the tangent plane: $dw = 2x\,dx + 2y\,dy - 2z\,dz = 4dx + 2dy - 2dz$. So $\Delta w \approx 4\Delta x + 2\Delta y - 2\Delta z$. The level surface is $\Delta w = 0$, its tangent plane approximation is $4\Delta x + 2\Delta y - 2\Delta z = 0$, i.e. $4(x-2)+2(y-1)-2(z-1)=0$, same as above).

Directional derivative. Rate of change of w as we move (x,y) in an arbitrary direction.

Take a unit vector $\hat u = \langle a,b\rangle$, and look at straight line trajectory $\vec r(s)$ with velocity $\hat u$, given by $x(s) = x_0 + as$, $y(s) = y_0 + bs$. (unit speed, so s is arclength!)

Notation: $\dfrac{dw}{ds}_{|\hat u}$.

Geometrically: slice of graph by a vertical plane (not parallel to x or y axes anymore). Directional derivative is the slope. Shown on applet (Functions of two variables), with $w = x^2+y^2+1$, and rotating slices through a point of the graph.

Know how to calculate dw/ds by chain rule: $\dfrac{dw}{ds}_{|\hat u} = \nabla w \cdot \dfrac{d\vec r}{ds} = \nabla w \cdot \hat u$.

Geometric interpretation: $dw/ds = \nabla w \cdot \hat u = |\nabla w|\cos\theta$. Maximal for $\cos\theta = 1$, when $\hat u$ is in direction of ∇w. Hence: direction of ∇w is that of fastest increase of w, and $|\nabla w|$ is the directional derivative in that direction. We have $dw/ds = 0$ when $\hat u \perp \nabla w$, i.e. when $\hat u$ is tangent to direction of level surface.

18.02 Lecture 13. – **Fri, Oct 5, 2007 (estimated – written before lecture)**

Practice exams 2A and 2B are on course web page.

Lagrange multipliers.

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

Example: find point of $xy=3$ closest to origin ? I.e. minimize $\sqrt{x^2+y^2}$, or better $f(x,y) = x^2+y^2$, subject to $g(x,y) = xy = 3$. Illustrated using Lagrange multipliers applet.

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 equations involving 3 variables x, y, λ:

$$\begin{cases} f_x = \lambda g_x \\ f_y = \lambda g_y \\ g = c \end{cases} \qquad \text{i.e. here} \qquad \begin{cases} 2x = \lambda y \\ 2y = \lambda x \\ xy = 3. \end{cases}$$

In general solving may be hard and require a computer. Here, linear algebra: $\begin{cases} 2x - \lambda y = 0 \\ -\lambda x + 2y = 0 \end{cases}$ requires either $x = y = 0$ (impossible, since $xy = 3$), or $\det = 4 - \lambda^2 = 0$. So $\lambda = \pm 2$. No solutions for $\lambda = -2$, while $\lambda = 2$ gives $(\sqrt{3}, \sqrt{3})$ and $(-\sqrt{3}, -\sqrt{3})$. (Checked on applet that $\nabla f = 2\nabla g$ at minimum).

Why the method works: at constrained min/max, moving in any direction along the constraint surface $g = c$ should give $df/ds = 0$. So, for any $\hat{u}$ tangent to $\{g = c\}$, $\frac{df}{ds}_{|\hat{u}} = \nabla f \cdot \hat{u} = 0$, i.e. $\hat{u} \perp \nabla f$. Therefore ∇f is normal to tangent plane to $g = c$, and so is ∇g, hence the gradient vectors are parallel.

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

Advanced example: surface-minimizing pyramid.

Triangular-based pyramid with given triangle as base and given volume V, using as little surface area as possible.

Note: $V = \frac{1}{3} A_{base}\, h$, so height h is fixed, top vertex moves in a plane $z = h$.

We can set up problem in coordinates: base vertices $P_1 = (x_1, y_1, 0)$, P_2, P_3, and top vertex $P = (x, y, h)$. Then areas of faces $= \frac{1}{2}|\vec{PP_1} \times \vec{PP_2}|$, etc. Calculations to find critical point of function of (x, y) are very hard.

Key idea: use variables adapted to the geometry, instead of (x, y): let a_1, a_2, a_3 = lengths of sides of the base triangle; u_1, u_2, u_3 = distances in the xy-plane from the projection of P to the sides of the base triangle. Then each face is a triangle with base length a_i and height $\sqrt{u_i^2 + h^2}$ (using Pythagorean theorem).

So we must minimize $f(u_1, u_2, u_3) = \frac{1}{2} a_1 \sqrt{u_1^2 + h^2} + \frac{1}{2} a_2 \sqrt{u_2^2 + h^2} + \frac{1}{2} a_3 \sqrt{u_3^2 + h^2}$.

Constraint? (asked using flashcards; this was a bad choice, very few students responded at all.) Decomposing base into 3 smaller triangles with heights u_i, we must have $g(u_1, u_2, u_3) = \frac{1}{2} a_1 u_1 + \frac{1}{2} a_2 u_2 + \frac{1}{2} a_3 u_3 = A_{base}$.

Lagrange multiplier method: $\nabla f = \lambda \nabla g$ gives

$$\frac{a_1}{2} \frac{u_1}{\sqrt{u_1^2 + h^2}} = \lambda \frac{a_1}{2}, \quad \text{similarly for } u_2 \text{ and } u_3.$$

We conclude $\lambda = \dfrac{u_1}{\sqrt{u_1^2 + h^2}} = \dfrac{u_2}{\sqrt{u_2^2 + h^2}} = \dfrac{u_3}{\sqrt{u_3^2 + h^2}}$, hence $u_1 = u_2 = u_3$, so P lies above the incenter.