18.02 Lecture 8. **– Tue, Sept 25, 2007**

Functions of several variables.

Recall: for a function of 1 variable, we can plot its graph, and the derivative is the slope of the tangent line to the graph.

Plotting graphs of functions of 2 variables: examples $z = -y$, $z = 1 - x^2 - y^2$, using slices by the coordinate planes. (derived carefully).

Contour plot: level curves $f(x, y) = c$. Amounts to slicing the graph by horizontal planes $z = c$.

Showed 2 examples from "real life": a topographical map, and a temperature map, then did the examples $z = -y$ and $z = 1 - x^2 - y^2$. Showed more examples of computer plots ($z = x^2 + y^2$, $z = y^2 - x^2$, and another one).

Contour plot gives some qualitative info about how f varies when we change x, y. (shown an example where increasing x leads f to increase).

Partial derivatives.

$f_x = \dfrac{\partial f}{\partial x} = \lim\limits_{\Delta x \to 0} \dfrac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$; same for f_y.

Geometric interpretation: f_x, f_y are slopes of tangent lines of vertical slices of the graph of f (fixing $y = y_0$; fixing $x = x_0$).

How to compute: treat x as variable, y as constant.

Example: $f(x, y) = x^3 y + y^2$, then $f_x = 3x^2 y$, $f_y = x^3 + 2y$.

18.02 Lecture 9. **– Thu, Sept 27, 2007**

Handouts: PS3 solutions, PS4.

Linear approximation

Interpretation of f_x, f_y as slopes of *slices* of the graph by planes parallel to xz and yz planes.

Linear approximation formula: $\Delta f \approx f_x \Delta x + f_y \Delta y$.

Justification: f_x and f_y give slopes of two lines tangent to the graph:

$y = y_0,\ z = z_0 + f_x(x_0, y_0)(x - x_0)$ and $x = x_0,\ z = z_0 + f_y(x_0, y_0)(y - y_0)$.

We can use this to get the equation of the tangent plane to the graph:

$$z = z_0 + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Approximation formula = the graph is close to its tangent plane.

Min/max problems.

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 tangent plane is horizontal (approximation formula: $\Delta f \simeq 0$, or rather, $|\Delta f| \ll |\Delta x|, |\Delta y|$).

Def of critical 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.

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).

Is it a max, min or saddle? (pictures shown of each type). Systematic answer: next lecture.

For today: observe $f = (x-y)^2 + 2y^2 + 2x - 2y = (x-y+1)^2 + 2y^2 - 1 \geq -1$, so minimum.

Least squares.

Set up problem: 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 $D = \sum_i (y_i - (ax_i + b))^2$.

$\dfrac{\partial D}{\partial a} = 0$ and $\dfrac{\partial D}{\partial b} = 0$ leads to a 2×2 linear system for a and b (done in detail as in Notes LS):

$$\left(\sum x_i^2\right) a + \left(\sum x_i\right) b = \sum x_i y_i$$
$$\left(\sum x_i\right) a + nb = \sum y_i$$

Least-squares setup also works in other cases: e.g. exponential laws $y = ce^{ax}$ (taking logarithms: $\ln y = \ln c + ax$, 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: Moore's Law (number of transistors on a computer chip increases exponentially with time): showed interpolation line on a log plot.

18.02 Lecture 10. – **Fri, Sept 28, 2007**

Second derivative test.

Recall critical points can be local min $(w = x^2 + y^2)$, local max $(w = -x^2 - y^2)$, saddle $(w = y^2 - x^2)$; slides shown of each type.

Goal: determine type of a critical point, and find the global min/max.

Note: global min/max may be either at a critical point, or on the boundary of the domain/at infinity.

We start with the case of $w = ax^2 + bxy + cy^2$, at $(0, 0)$.

Example from Tuesday: $w = x^2 - 2xy + 3y^2$: completing the square, $w = (x-y)^2 + 2y^2$, minimum.

If $a \neq 0$, then $w = a(x^2 + \frac{b}{a}xy) + cy^2 = a(x + \frac{b}{2a}y)^2 + (c - \frac{b^2}{4a})y^2 = \frac{1}{4a}(4a^2(x + \frac{b}{2a}y)^2 + (4ac - b^2)y^2)$.

3 cases: if $4ac - b^2 > 0$, same signs, if $a > 0$ then minimum, if $a < 0$ then maximum; if $4ac - b^2 < 0$, opposite signs, saddle; if $4ac - b^2 = 0$, degenerate case.

This is related to the quadratic formula: $w = y^2\big(a(\frac{x}{y})^2 + b(\frac{x}{y}) + c\big)$.

If $b^2 - 4ac < 0$ then no roots, so $at^2 + bt + c$ has a constant sign, and w is either always nonnegative or always nonpositive (min or max). If $b^2 - 4ac > 0$ then $at^2 + bt + c$ crosses zero and changes sign, so w can have both signs, saddle.

General case: second derivative test.

We look at second derivatives: $f_{xx} = \dfrac{\partial^2 f}{\partial x^2}$, f_{xy}, f_{yx}, f_{yy}. Fact: $f_{xy} = f_{yx}$.

Given f and a critical point (x_0, y_0), set $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$, $C = f_{yy}(x_0, y_0)$, then:

– if $AC - B^2 > 0$ then: if $A > 0$ (or C), local min; if $A < 0$, local max.

– if $AC - B^2 < 0$ then saddle.

– if $AC - B^2 = 0$ then can't conclude.

Checked quadratic case ($f_{xx} = 2a = A$, $f_{xy} = b = B$, $f_{yy} = 2c = C$, then $AC - B^2 = 4ac - b^2$).

General justification: quadratic approximation formula (Taylor series at order 2):

$$\Delta f \simeq f_x\,(x - x_0) + f_y\,(y - y_0) + \frac{1}{2} f_{xx}\,(x - x_0)^2 + f_{xy}\,(x - x_0)(y - y_0) + \frac{1}{2} f_{yy}\,(y - y_0)^2.$$

At a critical point, $\Delta f \simeq \frac{A}{2}(x - x_0)^2 + B(x - x_0)(y - y_0) + \frac{C}{2}(y - y_0)^2$. In degenerate case, would need higher order derivatives to conclude.

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

Example: $f(x, y) = x + y + \frac{1}{xy}$, for $x > 0$, $y > 0$.

$f_x = 1 - \frac{1}{x^2 y} = 0$, $f_y = 1 - \frac{1}{xy^2} = 0$. So $x^2 y = 1$, $xy^2 = 1$, only critical point is $(1, 1)$.

$f_{xx} = 2/x^3 y$, $f_{xy} = 1/x^2 y^2$, $f_{yy} = 2/xy^3$. So $A = 2$, $B = 1$, $C = 2$.

Question: type of critical point? Answer: $AC - B^2 = 2 \cdot 2 - 1 > 0$, $A = 2 > 0$, local min.

What about the maximum? Answer: $f \to \infty$ near boundary ($x \to 0$ or $y \to 0$) and at infinity.

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.

18.02 Lecture 14. – Thu, Oct 11, 2007

Handouts: PS5 solutions, PS6, practice exams 2A and 2B.

Non-independent variables.

Often we have to deal with non-independent variables, e.g. $f(P, V, T)$ where $PV = nRT$.

Question: if $g(x, y, z) = c$ then can think of $z = z(x, y)$. What are $\partial z/\partial x$, $\partial z/\partial y$?

Example: $x^2 + yz + z^3 = 8$ at $(2, 3, 1)$. Take differential: $2x\,dx + z\,dy + (y + 3z^2)\,dz = 0$, i.e. $4\,dx + dy + 6\,dz = 0$ (constraint $g = c$), or $dz = -\frac{4}{6}\,dx - \frac{1}{6}\,dy$. So $\partial z/\partial x = -4/6 = -2/3$ and $\partial z/\partial y = -1/6$ (taking the coefficients of dx and dy). Or equivalently: if y is held constant then we substitute $dy = 0$ to get $dz = -4/6\,dx$, so $\partial z/\partial x = -4/6 = -2/3$.

In general: $g(x, y, z) = c \Rightarrow g_x\,dx + g_y\,dy + g_z\,dz = 0$. If y held fixed, get $g_x\,dx + g_z\,dz = 0$, i.e. $dz = -g_x/g_z\,dx$, and $\partial z/\partial x = -g_x/g_z$.

Warning: notation can be dangerous! For example:

$f(x, y) = x + y$, $\partial f/\partial x = 1$. Change of variables $x = u$, $y = u + v$ then $f = 2u + v$, $\partial f/\partial u = 2$.

$x = u$ but $\partial f/\partial x \neq \partial f/\partial u$!!

This is because $\partial f/\partial x$ means change x keeping y fixed, while $\partial f/\partial u$ means change u keeping v fixed, i.e. change x keeping $y - x$ fixed.

When there's ambiguity, we must precise what is held fixed: $\left(\dfrac{\partial f}{\partial x}\right)_y$ = deriv. / x with y held fixed, $\left(\dfrac{\partial f}{\partial u}\right)_v$ = deriv. / u with v held fixed.

We now have $\left(\dfrac{\partial f}{\partial u}\right)_v = \left(\dfrac{\partial f}{\partial x}\right)_v \neq \left(\dfrac{\partial f}{\partial x}\right)_y$.

In above example, we computed $(\partial z/\partial x)_y$. When there is no risk of confusion we keep the old notation, by default $\partial/\partial x$ means we keep y fixed.

Example: area of a triangle with 2 sides a and b making an angle θ is $A = \frac{1}{2}ab\sin\theta$. Suppose it's a right triangle with b the hypothenuse, then constraint $a = b\cos\theta$.

3 ways in which rate of change of A w.r.t. θ makes sense:

1) view $A = A(a, b, \theta)$ independent variables, usual $\dfrac{\partial A}{\partial \theta} = A_\theta$ (with a and b held fixed). This answers the question: a and b fixed, θ changes, triangle stops being a right triangle, what happens to A?

2) constraint $a = b\cos\theta$, keep a fixed, change θ, while b does what it must to satisfy the constraint: $\left(\dfrac{\partial A}{\partial \theta}\right)_a$.

3) constraint $a = b\cos\theta$, keep b fixed, change θ, while a does what it must to satisfy the constraint: $\left(\dfrac{\partial A}{\partial \theta}\right)_b$.

How to compute e.g. $(\partial A/\partial\theta)_a$? [treat A as function of a and θ, while $b = b(a, \theta)$.]

0) Substitution: $a = b\cos\theta$ so $b = a\sec\theta$, $A = \frac{1}{2}ab\sin\theta = \frac{1}{2}a^2\tan\theta$, $(\frac{\partial A}{\partial\theta})_a = \frac{1}{2}a^2\sec^2\theta$. (Easiest here, but it's not always possible to solve for b)

1) Total differentials: $da = 0$ (a fixed), $dA = A_\theta d\theta + A_a da + A_b db = \frac{1}{2}ab\cos\theta\,d\theta + \frac{1}{2}b\sin\theta\,da + \frac{1}{2}a\sin\theta\,db$, and constraint $\Rightarrow da = \cos\theta\,db - b\sin\theta\,d\theta$. Plugging in $da = 0$, we get $db = b\tan\theta\,d\theta$

and then

$$dA = (\frac{1}{2}\,ab\cos\theta + \frac{1}{2}a\sin\theta\, b\tan\theta)d\theta, \quad \left(\frac{\partial A}{\partial\theta}\right)_a = \frac{1}{2}ab\cos\theta + \frac{1}{2}a\sin\theta\, b\tan\theta = \frac{1}{2}ab\sec\theta.$$

2) Chain rule: $(\partial A/\partial\theta)_a = A_\theta(\partial\theta/\partial\theta)_a + A_a(\partial a/\partial\theta)_a + A_b(\partial b/\partial\theta)_b = A_\theta + A_b(\partial b/\partial\theta)_a$. We find $(\partial b/\partial\theta)_a$ by using the constraint equation. [Ran out of time here]. Implicit differentiation of constraint $a = b\cos\theta$: we have $0 = (\partial a/\partial\theta)_a = (\partial b/\partial\theta)_a\cos\theta - b\sin\theta$, so $(\partial b/\partial\theta)_a = b\tan\theta$, and hence

$$\left(\frac{\partial A}{\partial\theta}\right)_a = \frac{1}{2}ab\cos\theta + \frac{1}{2}a\sin\theta\, b\tan\theta = \frac{1}{2}ab\sec\theta.$$

The two systematic methods essentially involve calculating the same quantities, even though things are written differently.

18.02 Lecture 15. – **Fri, Oct 12, 2007**

Review topics.

– Functions of several variables, contour plots.

– Partial derivatives, gradient; approximation formulas, tangent planes, directional derivatives.

Note: *partial differential equations* (= equations involving partial derivatives of an unknown function) are very important in physics. E.g., heat equation: $\partial f/\partial t = k(\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2 + \partial^2 f/\partial z^2)$ describes evolution of temperature over time.

– Min/max problems: critical points, 2nd derivative test, checking boundary.
(least squares won't be on the exam)

– Differentials, chain rule, change of variables.

– Non-independent variables: Lagrange multipliers, and constrained partial derivatives.

Re-explanation of how to compute constrained partials: say $f = f(x,y,z)$ where $g(x,y,z) = c$. To find $(\partial f/\partial z)_y$:

1) using differentials: $df = f_x\,dx + f_y\,dy + f_z\,dz$. We set $dy = 0$ since y held constant, and want to eliminate dx. For this we use the constraint: $dg = g_x\,dx + g_y\,dy + g_z\,dz = 0$, so setting $dy = 0$ we get $dx = -g_z/g_x\,dz$. Plug into df: $df = -f_x g_z/g_x\,dz + g_z\,dz$, so $(\partial f/\partial z)_y = -f_x g_z/g_x + g_z$.

2) using chain rule: $\left(\frac{\partial f}{\partial z}\right)_y = \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial z}\right)_y + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial z}\right)_y + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial z}\right)_y = f_x\left(\frac{\partial x}{\partial z}\right)_y + f_z$, while

$$0 = \left(\frac{\partial g}{\partial z}\right)_y = \frac{\partial g}{\partial x}\left(\frac{\partial x}{\partial z}\right)_y + \frac{\partial g}{\partial y}\left(\frac{\partial y}{\partial z}\right)_y + \frac{\partial g}{\partial z}\left(\frac{\partial z}{\partial z}\right)_y = g_x\left(\frac{\partial x}{\partial z}\right)_y + g_z$$

which gives $(\partial x/\partial z)_y$ and hence the answer.