REVIEW SHEET FOR FINAL: BASIC

MATH 195, SECTION 59 (VIPUL NAIK)

The document does not include material that was part of the midterm 1 and midterm 2 review sessions. Please also bring copies of these review sheets to the review session on Monday.

You are expected to review this on your own time. We will concentrate on the *advanced* review sheet during problem session.

1. Directional derivatives and gradient vectors

Words ...

(1) The directional derivative of a scalar function f of two variables along a *unit* vector $\mathbf{u} = a\mathbf{i} + b\mathbf{j}$ at a point (x_0, y_0) is defined as the following limit of difference quotient, if the limit exists:

$$\lim_{h \to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h}$$

(2) The directional derivative of a differentiable scalar function f of two variables along a *unit* vector $\mathbf{u} = a\mathbf{i} + b\mathbf{j}$ at a point (x_0, y_0) is $D_{\mathbf{u}}(f) = af_x(x_0, y_0) + bf_y(x_0, y_0)$.

(3) The gradient vector for a *differentiable* scalar function f of two variables at a point (x_0, y_0) is $\nabla f(x_0, y_0) = f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j}$.

(4) The directional derivative of f is the dot product of the gradient vector of ∇f and the unit vector $\mathbf{u}$.

(5) Suppose ∇f is nonzero. Then, if $\mathbf{u}$ makes an angle θ with ∇f, then $D_{\mathbf{u}}(f)$ is $|\nabla_f| \cos \theta$. The directional derivative is maximum in the direction of the gradient vector, zero in directions orthogonal to the gradient vector, and minimum in the direction opposite to the gradient vector.

(6) The level curves are orthogonal to the gradient vector.

(7) Similar formulas for gradient vector and directional derivative work in three dimensions.

(8) The level surfaces are orthogonal to the gradient vector for a function of three variables.

(9) For a surface given by $F(x, y, z) = 0$, if (x_0, y_0, z_0) is a point on the surface, and $F_x(x_0, y_0, z_0)$, $F_y(x_0, y_0, z_0)$, and $F_z(x_0, y_0, z_0)$ all exist and are nonzero, then the normal line is given by:

$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}$$

The tangent plane is given by:

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

2. Max-min values

Words ...

(1) For a directional local minimum, the directional derivative (in the outward direction from the point) is greater than or equal to zero. For a directional local maximum, the directional derivative (in the outward direction from the point) is less than or equal to zero.

Note that even for *strict* directional local maximum or minimum, the possibility of the directional derivative being zero cannot be ruled out.

(2) If a point is a point of directional local minimum from two opposite directions (i.e., it is a local minimum along a line through the point, from both directions on the line) then the directional derivative along the line, if it exists, must equal zero.

(3) If a function of two variables is differentiable at a point of local minimum or local maximum, then the directional derivative of the function is zero at the point in every direction. Equivalently, the gradient vector of the function at the point is the zero vector. Equivalently, both the first partial derivatives at the point are zero.

Points where the gradient vector is zero are termed *critical points*.

(4) If the directional derivatives along some directions are positive and the directional derivatives along other directions are negative, the point is likely to be a *saddle point*. A saddle point is a point for which the tangent plane to the surface that's the graph of the function slides through the graph, i.e., it is not completely on one side.

(5) For a function f of two variables with continuous second partials, and a critical point (a, b) in the domain (so $f_x(a, b) = f_y(a, b) = 0$) we compute the Hessian determinant:

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

If $D(a, b) > 0$ and $f_{xx}(a, b) > 0$, the function has a local *minimum* at the point (a, b). If $D(a, b) > 0$ and $f_{xx}(a, b) < 0$, the function has a local *maximum* at the point (a, b). If $D(a, b) < 0$, we get a saddle point at the point. If $D(a, b) = 0$, the situation is inconclusive, i.e., the test is indecisive.

(6) For a closed bounded subset of $\mathbb{R}^n$ (and specifically $\mathbb{R}^2$) any continuous function with domain that subset attains its absolute maximum and minimum values. These values are attained either at interior points (in which case they are local extreme values and must be attained at critical points) or at boundary points.

(7) *Relation with level curves*: *Typically*, local extreme values correspond to isolated single point level curves. However, this is not always the case, and there are some counterexamples. To be more precise, any *isolated* or *strict* local extreme value corresponds to a (locally) single point level curve.

Actions ...

(1) Strategy for finding local extreme values: First, find all the critical points by solving $f_x(a, b) = 0$ and $f_y(a, b) = 0$ as a pair of simultaneous equations. Next, use the second derivative test for each critical point, and if feasible, try to figure out if this is a point of local maximum, or local minimum, or a saddle point.

(2) To find absolute extreme values of a function on a closed bounded subset of $\mathbb{R}^2$, first find critical points, then find critical points for a parameterization of the boundary, and then compute values at all of these and see which is largest and smallest. *If the list of critical points is finite, and we need to find absolute maximum and minimum, it is not necessary to do the second derivative test to figure out which points give local maximum, local minimum, or neither, we just need to evaluate at all points and find the maximum/minimum.*

(3) When the domain of the function is bounded but not closed, we must consider the possibility of extreme values occurring as we approach boundary points not in the domain. If the domain is not bounded, we must consider directions of approach to infinity.

3. Lagrange multipliers

Words ...

(1) Two of the reasons why the derivative of a function may be zero: the function is constant around the point, or the function has a local extreme value at the point.

Version for many variables: two of the reasons why the gradient vector of a function of many variables may be zero: the function is constant around the point, or the function has a local extreme value at the point.

Version for function restricted to a subset smooth around a point: two of the reasons why the gradient vector may be *orthogonal* to the subset at the point: the function is constant on the subset around the point, or the function has a local extreme value (relative to the subset) at the point.

(2) For a function f defined on a subset smooth around a point (i.e., with a well defined tangent and normal space), if f has a local extreme value at the point when restricted to the subset, then ∇f lives in the normal direction to the subset (this includes the possibility of it being zero).

(3) For a codimension one subset of $\mathbb{R}^n$ defined by a condition $g(x_1, x_2, \dots, x_n) = k$, if a point $(a_1, a_2, \dots, a_n)$ gives a local extreme value for a function f of n variables, and if ∇g is well defined and nonzero at the point, then there exists a real number λ such that $\nabla f(a_1, a_2, \dots, a_n) = \lambda \nabla g(a_1, a_2, \dots, a_n)$. Note that λ may be zero.

(4) Suppose a codimension r subset of $\mathbb{R}^n$ is given by r independent constraints $g_1(x_1, x_2, \dots, x_n) = k_1$, $g_2(x_1, x_2, \dots, x_n) = k_2$, and so on till $g_r(x_1, x_2, \dots, x_n) = k_r$. Suppose ∇g_i are nonzero for all i at a point $(a_1, a_2, \dots, a_n)$ of local extreme value for a function f relative to this subset. Suppose further that all the ∇g_i are linearly independent. Then $\nabla f(a_1, a_2, \dots, a_n)$ is a linear combination of the vectors $\nabla g_1(a_1, a_2, \dots, a_n)$, $\nabla g_2(a_1, a_2, \dots, a_n)$, $\dots$, $\nabla g_r(a_1, a_2, \dots, a_n)$. In other words, there exist real numbers λ_1, λ_2, $\dots$, λ_r such that:

$$\nabla f(a_1, a_2, \dots, a_n) = \lambda_1 \nabla g_1(a_1, a_2, \dots, a_n) + \lambda_2 \nabla g_2(a_1, a_2, \dots, a_n) + \cdots + \lambda_r \nabla g_r(a_1, a_2, \dots, a_n)$$

(5) The Lagrange condition may be violated at points of local extremum where ∇g is zero, or more generally, where the ∇g_i fail to be linearly independent. This may occur either because the tangent and normal space are not well defined or because the functions fail to capture it well.

Actions ...

(1) Suppose we want to maximize and minimize f on the set $g(x_1, x_2, \dots, x_n) = k$. Assume $\nabla g(x_1, x_2, \dots, x_n)$ is defined everywhere on the set and never zero. Suppose ∇f is also defined. Then, all local maxima and local minima are attained at points where $\nabla f = \lambda \nabla g$ for some real number λ. To find these, we solve the system of $n+1$ equations in the $n+1$ variables x_1, x_2, $\dots$, x_n, namely the n scalar equations from the Lagrange condition and the equation $g(x_1, x_2, \dots, x_n) = k$.

    To find the actual extreme values, once we've collected all candidate points from the above procedure, we evaluate the function at all these and find the largest and smallest value to find the absolute maximum and minimum.

(2) If there are points in the domain where ∇g takes the value 0, these may also be candidates for local extreme values, and the function should additionally be evaluated at these as well to find the absolute maximum and minimum.

(3) A similar procedure works for a subset given by r constraints. In this case, we have the equation:

$$\nabla f(a_1, a_2, \dots, a_n) = \lambda_1 \nabla g_1(a_1, a_2, \dots, a_n) + \lambda_2 \nabla g_2(a_1, a_2, \dots, a_n) + \cdots + \lambda_r \nabla g_r(a_1, a_2, \dots, a_n)$$

as well as the r equations $g_1(x_1, x_2, \dots, x_n) = k_1$, $g_2(x_1, x_2, \dots, x_n) = k_2$, and so on. In total, we have $n+r$ equations in $n+r$ variables: the x_1, x_2, $\dots$, x_n and the λ_1, λ_2, $\dots$, λ_r.

4. Max-min values: examples

(1) *Additively separable, critical points*: For an additively separable function $F(x, y) := f(x) + g(y)$, the critical points of F are the points whose x-coordinate gives a critical point for f and y-coordinate gives a critical point for g.

(2) *Additively separable, local extreme values*: The local maxima occur at points whose x-coordinate gives a local maximum for f and y-coordinates gives a local maximum for g. Similarly for local minima. If one coordinate gives a local maximum and the other coordinate gives a local minimum, we get a saddle point.

(3) *Additively separable, absolute extreme values*: If the domain is a rectangular region, rectangular strip, or the whole plane, then the absolute maximum occurs at the point for which each coordinate gives the absolute maximum for that coordinate, and analogously for absolute minimum. This does *not* work for non-rectangular regions in general.

(4) *Multiplicatively separable, critical points*: For a multiplicatively separable function $F(x, y) := f(x)g(y)$ with f, g, differentiable, there are four kinds of critical points (x_0, y_0): (1) $f'(x_0) = g'(y_0) = 0$, (2) $f(x_0) = f'(x_0) = 0$, (3) $g(y_0) = g'(y_0) = 0$, (4) $f(x_0) = g(y_0) = 0$.

(5) *Multiplicatively separable, local extreme values*: At a critical point of Type (1), the nature of local extreme value for F depends on the signs of f and g *and* on the nature of local extreme values for

each. See the table. Critical points of Type (4) alone do not give local extreme values. The situation with critical points of Types (2) and (3) is more ambiguous and too complicated for discussion.

(6) *Multiplicatively separable, absolute extreme values*: Often, these don't exist, if one function takes arbitrarily large magnitude values and the other one takes nonzero values (details based on sign). If both functions are everywhere positive, and we are on a rectangular region, then the absolute maximum/minimum for the product occur at points whose coordinates give respective absolute maximum/minimum for f and g. (See notes)

(7) For a continuous quasiconvex function on a convex domain, the maximum must occur at one of the extreme points, in particular on the boundary. If the function is strictly quasiconvex, the maximum can occur only at a boundary point.

(8) For a continuous quasiconvex function on a convex domain, the minimum must occur on a convex subset. If the function is strictly quasiconvex, it must occur at a unique point.

(9) Linear functions are quasiconvex but not strictly so. The negative of a linear function is also quasiconvex. The maximum and minimum for linear functions on convex domains must occur at extreme points.

(10) To find maxima/minima on the boundary, we can use the method of Lagrange multipliers.

See also: tables, discussion for linear, quadratic, and homogeneous functions (hard to summarize). Below is a copy of the table for the multiplicatively separable case.

The setup here is that we have a function $F(x, y) := f(x)g(y)$ and a point (x_0, y_0) in the domain such x_0 is a critical point for f and y_0 is a critical point for g. Visit the lecture notes for more detailed context.

$f(x_0)$ sign	$g(y_0)$ sign	$f(x_0)$ (local max/min)	$g(y_0)$ (local max/min)	$F(x_0, y_0)$ (local max/min/saddle)
positive	positive	local max	local max	local max
positive	positive	local max	local min	saddle point
positive	positive	local min	local max	saddle point
positive	positive	local min	local min	local min
positive	negative	local max	local max	saddle point
positive	negative	local max	local min	local min
positive	negative	local min	local max	local max
positive	negative	local min	local min	saddle point
negative	positive	local max	local max	saddle point
negative	positive	local max	local min	local max
negative	positive	local min	local max	local min
negative	positive	local min	local min	saddle point
negative	negative	local max	local max	local min
negative	negative	local max	local min	saddle point
negative	negative	local min	local max	saddle point
negative	negative	local min	local min	local max