Extreme Values of Functions of Several Variables

Math 214 Handout

October 22, 2003

Recall that if S is a subset of n-dimensional space and P is a point of S we say that P is a point in the *interior* of S or a point *inside* S if there is some (small) positive number r such that every point of n-dimensional space within distance r of P is a point of S.

Recall that a function f of n variables is *differentiable* at a point inside its domain if it admits first order approximation by a linear function near the given point.

Theorem: If a function f of n variables has an extreme value for the subset S of its domain at a point P of S that is a point *inside* the domain of f where f is differentiable, then the gradient vector $\nabla f(P)$ of f at P must be perpendicular to the tangent vector at P of every differentiably parameterized curve lying in S and passing through P.

Proof. Let $G(t)$ be a differentiably parameterized curve contained in S and passing through P when $t = a$. Since S is contained in the domain of f, the function $h(t) = f(G(t))$ is defined for all values of t for which $G(t)$ is defined, and since f is differentiable at $P = G(a)$, the function h is differentiable at a. In fact, the "chain rule" tells us that

$$h'(a) \;=\; \nabla f(P) \cdot G'(a) \quad .$$

Since f has an extreme value relative to the set S at the point P and each $G(t)$ is in S, it follows that h, a function of one variable, has a local extreme value at $t = a$, and, therefore, that $h'(a) = 0$. Consequently, $\nabla f(P)$ is perpendicular to the tangent vector $G'(a)$ of the curve at P.

Corollary 1. If a function f of n variables has an extreme value for the subset S of its domain at a point P of S that is a point *inside* S where f is differentiable, then the gradient vector $\nabla f(P)$ must be the zero vector.

Proof. If P is a point *inside* S then every sufficiently short line segment passing through P must be perpendicular to $\nabla f(P)$, which means that every vector must be perpendicular to $\nabla f(P)$.

Corollary 2. If a function f of n variables has an extreme value for the subset $S = \{g = 0\}$ of its domain at a point P of S where f and g are differentiable functions, then the gradient $\nabla f(P)$ of f and the gradient $\nabla g(P)$ of g must be parallel vectors.

Proof. The statement is formally true, but probably useless if $\nabla g(P) = 0$. We assume that $\nabla g(P)$ is not the zero vector. In this case ∇g is perpendicular to the tangent hyperplane (i.e., plane if $n = 3$ or line if $n = 2$) to S at P. Every unit vector in the tangent hyperplane is tangent to some small differentiably parameterized curve segment lying in S and passing through P. Hence, by the theorem, $\nabla f(P)$ is also perpendicular to each such curve segment, and, hence, to the tangent hyperplane. Since a hyperplane has only one parallel class of normal vectors, $\nabla f(P)$ and $\nabla g(P)$ must be parallel.

Remark. The theorem is useful also in the case where f is a function of 3 variables and the constraint set S is a curve in space. Then the fact that P lies in S corresponds roughly to two equations for P and the orthogonality condition of the theorem provides, in non-degenerate situations an additional equation with the result that (usually) only finitely many such P are possible. (Among these are points that are maxima, minima, and those that are neither.) This is equivalent to the principle of "Lagrange multipliers" discussed in the text.

Document network location for HTML:

`http://math.albany.edu:8000/math/pers/hammond/course/mat214/extremes.html`