Critical Points

A point P is a **critical point** of f if $\operatorname{grad} f(P) = O$. Equivalently, all the partial derivatives $D_i f$ are 0 at P.

Example. Find the critical points of $f(x, y) = e^{-(x^2+y^2)}$. We take partial derivatives and set them to 0 to find the critical points.

As in the single variable case, we can have a variety of behaviors at a critical point; we do not necessarily have a local minimum or local maximum.

Let f be defined on an open set U. A point P is called a **local maximum** of f if, in some neighborhood N of P, we have

$$f(X) \leq f(P)$$

for all $X \in N$.

The concept of local minimum is defined similarly.

Theorem. Let f be a differentiable function on U. Let P be a local maximum. Then P is a critical point of f.

The proof of this amounts to reducing it to a one variable problem. If H is a nonzero vector, and t is small enough, then $P + tH \in U$. Moreover, if t is small enough, $P + tH$ will land in the neighborhood mentioned in the definition, so that

$$f(P + tH) \leq f(P)$$

for all t in an interval of the form $(-\delta, \delta)$, $\delta > 0$. So $g(t) = f(P + tH)$ has a local maximum at $t = 0$. Thus $g'(t) = 0$. By the chain rule,

$$\operatorname{grad} f(P) \cdot H = 0.$$

This is true for all H, so we must have $\operatorname{grad} f(P) = 0$. ■