1 Recap

- Clairaut's theorem.
- Taylor's theorem and extrema in one-variable.
- First derivative test in more than one variable.

2 Extrema in more than one variable

Find the global extrema of $f(x,y,z) = x^2 - y^2 + 3z^2$ on $x^2 + y^2 + z^2 \leq 1$.
f is diff everywhere. Let's look at critical points first: $\nabla f = (2x, -2y, 6z)$ which vanishes only at the origin (which lies in S). The value of f there is 0. On the boundary of S, i.e, on the sphere $x^2+y^2+z^2 = 1$, We see that $f(x,y) = x^2-y^2+3(1-x^2-y^2) = 3-2x^2-4y^2$ on $x^2+y^2 \leq 1$. Now again let's look at critical points: $\nabla f = (-4x, -8y)$ which is 0 at $(0,0)$ lying in $x^2+y^2 \leq 1$. The value of f is 3 there. Let's look at the boundary $x^2+y^2 = 1$. There, $f(x) = 3-2x^2-4(1-x^2) = -1+2x^2$ and $-1 \leq x \leq 1$. Again $f' = 4x = 0$ when $x = 0 \in [-1,1]$. There $f(0) = -1$. At the end-points, $f(-1) = f(1) = 1$. Thus the global max value is 3 occuring at $(0,0,\pm 1)$ and the global min value is -1 occuring at $(0,\pm 1, 0)$.

Before formulating a second-derivative test for local extrema, note this curious phenomenon: Consider $f(x,y) = x^2 - y^2$. Note that $\nabla f = (2x, -2y) = (0,0)$ when $(x,y) = (0,0)$. Note that f does *not* assume a local extremum at $(0,0)$. This is *not* because the second derivatives vanish. Indeed, $f_{xx} = 2, f_{yy} = -2, f_{xy} = f_{yx} = 0$. Rather, in some direction(s) that is, along $(0,1)$, f decreases and in some other(s) (along $(1,0)$) it increases.
Definition: A critical point is said to be a *saddle* point if *every* open ball containing $\vec{a}$ lying completely in the domain, contains points $\vec{r}_1, \vec{r}_2$ such that $f(\vec{r}_1) > f(\vec{a})$ and $f(\vec{r}_2) < f(\vec{a})$. In the example above $(0,0)$ is a saddle point.

Let $\vec{a}$ be a critical point of f. Suppose f is C^3 in a neighbourhood of $\vec{a}$ (that is, the first, second, and third partials exist in a neighbourhood of a and are continuous there; by Clairut, the mixed partials are equal).
Theorem: Under the above assumptions, for all $\vec{h}$ lying in a certain neighbourhood of $\vec{0}$, $|f(\vec{a}+\vec{h}) - f(\vec{a}) - \nabla_{\vec{h}} f(\vec{a}) - \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) h_i h_j| \leq C\|h\|^3$ for some $C > 0$.
Proof: Consider $u(t) = f(\vec{a}+t\vec{h})$. By any application of the chain rule and properties of continuity, we see that $u(t)$ is C^3 in $(-\epsilon, \epsilon)$ for some $\epsilon > 0$. Applying a precise version of the one-variable Taylor theorem, it turns out that $|u(t) - u(0) - u'(0)t - \frac{u''(0)}{2}t^2| \leq C|t^3|$ in a neighbourhood of $t = 0$ and C does not depend on $\vec{h}$. Now $u'(0) = \nabla_{\vec{h}} f(\vec{a})$. In fact, $u'(t) = \sum_i \frac{\partial f}{\partial x_i}(\vec{a}+t\vec{h})h_i$. Thus $u''(0) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) h_i h_j$. Now replace t with $|h|$ and h with $\frac{h}{|h|}$ to get the result. □

Let $\vec{a}$ be a critical point of a scalar field f that is C^3 in a neighbourhood of $\vec{a}$. Then if $\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) h_i h_j > 0$ for all $\vec{h} \neq \vec{0}$, i.e., the symmetric matrix $H(\vec{a})$ (the Hessian) given by $H_{ij}(\vec{a}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})$ is positive-definite, then $\vec{a}$ is a local minimum. If $H(\vec{a})$ is

negative-definite, then $\vec{a}$ is a local maximum. If H is invertible but neither positive nor negative definite, then $\vec{a}$ is a saddle point. (If H is not invertible, pray to the flying spaghetti monster.)
This result raises the question "How does one figure out if a Hermitian matrix H is positive-definite or not?"
Answer: Since H is Hermitian, it is diagonalisable as $H = U^\dagger DU$ for some unitary U. Thus $h^T H\bar{h} = h^T U^\dagger DU\bar{h} = \sum_i \lambda_i |(U\bar{h})_i|^2 = \sum_i |y_i|^2 \lambda_i$ where $y = U\bar{h}$. Thus this expression is positive for all h if and only if it is so for all y if and only $\lambda_i > 0$ for all i. Likewise, H is negative-definite if and only if all the eigenvalues are negative. It is invertible if and only if all of them are non-zero.

From the second-order Taylor expansion, $|f(\vec{a}+\vec{h}) - f(\vec{a}) - \frac{1}{2}h^T H(\vec{a})h| \leq C\|\vec{h}\|^3$. As above, diagonalise $H = O^D O$. Now $h^T H(\vec{a})h = \sum_i (Oh)_i^2 \lambda_i$. Thus, $\frac{1}{2}\sum_i |(Oh)_i|^2 \lambda_i - C\|\vec{h}\|^3 \leq f(\vec{a}+\vec{h}) - f(\vec{a}) \leq \frac{1}{2}\sum_i |(Oh)_i|^2 \lambda_i + C\|\vec{h}\|^3$. If $H(\vec{a})$ is positive-definite, then $\lambda_i > 0$. Let $\lambda_i > c > 0$ for all i. Thus, $f(\vec{a}+\vec{h}) - f(\vec{a}) \geq \frac{c}{2}\|\vec{h}\|^2 - C\|\vec{h}\|^3$. (Indeed, $\sum_i (Oh)_i^2 = h^T O^T Oh = h^T h = \|\vec{h}\|^2$.) If $\|h\| < \frac{c}{4C}$, then $f(\vec{a}+\vec{h}) \geq f(\vec{a})$. Thus it is a local min. Likewise for local max and saddle points (HW).

Find all local extrema of $f(x,y) = x^2 - xy - 2y^2$ on $x^2 + y^2 \leq 16$.
$\nabla f = (2x - y, -x - 4y) = (0,0)$ precisely when $(x,y) = (0,0)$. The second derivatives at $(0,0)$ are $f_{xx}(0,0) = 2$, $f_{yy}(0,0) = -4$, $f_{xy}(0,0) = f_{yx}(0,0) = -1$. Thus the Hessian matrix H is $\begin{bmatrix} 2 & -1 \\ -1 & -4 \end{bmatrix}$. Its eigenvalues can be computed to be $-1 \pm \sqrt{10}$ and hence it is a saddle point. That is, there are no local extrema in the region.
Ideally, we'd like to develop a method to handle local/global extrema when *constraints* are imposed. This method is called Lagrange's multipliers. However, we shall postpone/skip it for now.