Let $f(x,y) = 2xy - 2x^3y - 2xy^3$, the example from the demos.

1. Find the critical points

$$\nabla f(x,y) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle = \langle\, 2y - 6x^2y - 2y^3, 2x - 2x^3 - 6xy^2 \,\rangle$$

The critical points are points such that $\nabla f(x,y) = \vec{0}$ or, in our case,

$$2y - 6x^2y - 2y^3 = 0 \tag{1}$$

$$2x - 2x^3 - 6xy^2 = 0 \tag{2}$$

If $x = 0$, equation (2) is satisfied and equation (1) becomes $2y - 2y^3 = 0$ which has solutions $y = 0$, $y = \pm 1$.

If $y = 0$, equation (1) is satisfied and equation (2) becomes $2x - 2x^3 = 0$ which has solutions $x = 0$, $x = \pm 1$.

These give 5 distinct critical points, $(0,0)$, $(0,\pm 1)$ and $(\pm 1, 0)$.

If $xy \neq 0$, (1) and (2) become

$$1 - 3x^2 - y^2 = 0 \tag{3}$$

$$1 - x^2 - 3y^2 = 0 \tag{4}$$

Equation (4) - 3 Equation(3) gives the equation $-2 + 8x^2 = 0$ and so $x = \pm\frac{1}{2}$. For either value of x, $y = \pm\frac{1}{2}$ so there are 4 more distinct critical points $\left(\pm\frac{1}{2}, \pm\frac{1}{2}\right)$.

These nine points are all the critical points.

2. The second derivative

If the gradient is the first derivative then the gradient of the gradient should be the second derivative. Define the *Hessian* of f:

$$\mathcal{H}(f) = \begin{pmatrix} \nabla \dfrac{\partial f}{\partial x} \\ \nabla \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Assuming the higher partials are continuous, $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$ so the Hessian matrix is symmetric.

In our case $\nabla \dfrac{\partial f}{\partial x} = \langle -12xy, 2 - 6x^2 - 6y^2 \rangle$ and $\nabla \dfrac{\partial f}{\partial y} = \langle 2 - 6x^2 - 6y^2, -12xy \rangle$ so

$$\mathcal{H}(f) = \begin{pmatrix} -12xy & 2 - 6x^2 - 6y^2 \\ 2 - 6x^2 - 6y^2 & -12xy \end{pmatrix}$$

$$\mathcal{H}(f)(0,0) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \qquad \mathcal{H}(f)(\pm 1, 0) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix} \qquad \mathcal{H}(f)(0, \pm 1) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}$$

$$\mathcal{H}(f)(\tfrac{1}{2}, \tfrac{1}{2}) = \mathcal{H}(f)(-\tfrac{1}{2}, -\tfrac{1}{2}) = \begin{pmatrix} -3 & -1 \\ -1 & -3 \end{pmatrix} \qquad \mathcal{H}(f)(\tfrac{1}{2}, -\tfrac{1}{2}) = \mathcal{H}(f)(-\tfrac{1}{2}, \tfrac{1}{2}) = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

3. The second derivative test

The second derivative test says

(1) Compute the Hessian.
(2) Compute the determinant of the Hessian, D.
(3) (a) If $D = 0$ there is no test (you don't know what's going on.)
  (b) If $D < 0$ the point is a saddle point
  (c) If $D > 0$ then
    (i) if the diagonal entries are positive the point is a local minimum.
    (ii) if the diagonal entries are negative the point is a local maximum.

In (c) parts (i) and (ii), think about why the diagonal entries are either both positive or both negative.

In our example, the determinants of the Hessian for $(0,0)$ is -4, and for $(0,\pm 1)$ and $(\pm 1, 0)$ it is -16. Hence these five critical points are saddles.

For $(\pm\frac{1}{2}, \pm\frac{1}{2})$ the determinant is 8. For $\pm(\frac{1}{2}, \frac{1}{2})$ the diagonal entries are -3 so these critical points are local maxima. For $\pm(\frac{1}{2}, -\frac{1}{2})$ the diagonal entries are 3 so these critical points are local minima.

4. Basic examples

$x^2 + y^2$ clearly has a minimum at $(0,0)$: $\mathcal{H} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

$-(x^2 + y^2)$ clearly has a maximum at $(0,0)$: $\mathcal{H} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$.

$x^2 - y^2$ has a saddle point at $(0,0)$: $\mathcal{H} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$.

$x^4 + y^4$ clearly has a minimum at $(0,0)$: $\mathcal{H} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

$-(x^4 + y^4)$ has a minimum at $(0,0)$: $\mathcal{H} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

$x^4 - y^4$ has a saddle at $(0,0)$: $\mathcal{H} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

These last three examples all have the determinant of the Hessian equal to 0 and so demonstrate that when the determinant of the Hessian vanishes it is impossible to say anything without further thought.

5. Why does it work?

Suppose $\vec{a}$ is a critical point of $f(\vec{x})$ in the interior of the domain of f and that all partials exist and are continuous near $\vec{a}$. To study what sort of critical point we have, let $\vec{r}(t)$ be a curve such that $\vec{r}(0) = \vec{a}$. Let $g(t) = f\big(\vec{r}(t)\big)$. Since $g'(t) = \nabla f\big(\vec{r}(t)\big) \bullet \big(\vec{r}\,'(t)\big)$ and since $\vec{a}$ is a critical point, $g'(0) = 0$.

From fist year calculus, we can try to figure out what sort of critical point we have by computing $g''(0)$. Since $g'(t) = \nabla f\big(\vec{r}(t)\big) \bullet \big(\vec{r}\,'(t)\big)$,

$$g''(t) = \frac{d\nabla f\big(\vec{r}(t)\big)}{dt} \bullet \big(\vec{r}\,'(t)\big) + \nabla f\big(\vec{r}(t)\big) \bullet \big(\vec{r}\,''(t)\big)$$

The term $\dfrac{d\nabla f\big(\vec{r}(t)\big)}{dt}$ can be computed one coordinate at a time:

$$\frac{d\,\dfrac{\partial f\big(\vec{r}(t)\big)}{\partial x_i}}{dt} = \nabla\frac{\partial f\big(\vec{r}(t)\big)}{\partial x_i} \bullet \big(\vec{r}\,'(t)\big)$$

The Hessian $\mathcal{H}\big(f(\vec{r}(t)\big) = \begin{pmatrix} \nabla\dfrac{\partial f\big(\vec{r}(t)\big)}{\partial x_1} \\ \vdots \\ \nabla\dfrac{\partial f\big(\vec{r}(t)\big)}{\partial x_i} \\ \vdots \\ \nabla\dfrac{\partial f\big(\vec{r}(t)\big)}{\partial x_n} \end{pmatrix}$

Since we are at a critical point, $\nabla f\big(\vec{r}(t)\big) \bullet \big(\vec{r}\,''(t)\big) = 0$. In linear algebra you will learn how to write and work with the other term in the formula for $g''(0)$ but for now suppose $\vec{r}(0) = (a, b)$ and $\vec{r}\,'(0) = \langle A, B\rangle$. Then, at a critical point,

$$g''(0) = A^2\,\frac{\partial^2 f}{\partial x^2}(a,b) + 2AB\,\frac{\partial^2 f}{\partial x \partial y}(a,b) + B^2\,\frac{\partial^2 f}{\partial y^2}(a,b)$$

where

$$\mathcal{H}(a,b) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial y \partial x}(a,b) \\ \dfrac{\partial^2 f}{\partial x \partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{pmatrix}$$

To be a local maximum, $g''(0) < 0$ for all non-zero vectors $\langle A, B\rangle$; to be a local minimum, $g''(0) > 0$ for all non-zero vectors $\langle A, B\rangle$.

Are there any directions for which $g''(0) = 0$?

If $B = 0$, then $A \neq 0$ and so $g''(0) \neq 0$ unless $\dfrac{\partial^2 f}{\partial x^2}(a,b) = 0$, in which case the determinant is negative.

If $B \neq 0$ then $g''(0) = 0$ if and only if $\frac{A}{B}$ is a solution of

$$x^2\,\frac{\partial^2 f}{\partial x^2}(a,b) + 2x\,\frac{\partial^2 f}{\partial x \partial y}(a,b) + \frac{\partial^2 f}{\partial y^2}(a,b) = 0$$

We need to look at the discriminant

$$4\frac{\partial^2 f}{\partial x \partial y}(a,b)^2 - 4\frac{\partial^2 f}{\partial x^2}(a,b)\frac{\partial^2 f}{\partial x^y}(a,b) = -4D$$

where D is the formula in the book.

If $D < 0$ then the quadratic has roots. Near the root the quadratic will be positive on one side and negative on the other so in some directions we have a local maxima and in others a local minima.

If $D > 0$ the quadratic has no roots and if $\dfrac{\partial^2 f}{\partial x^2}(a,b) > 0$ the quadratic is always positive and hence $g''(0) > 0$ and all curves through the critical point have a local minimum there. If $D > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(a,b) < 0$ $g''(0) < 0$ and we have a local maximum.

6. Absolute Max-Min

A *global maximum value* is a number M such that $M \geqslant f(\vec{x}\,)$ for all $\vec{x} \in D$ where D is the domain. If f is continuous and if D is closed and bounded, then there exist one or more points $\vec{a}$ such that $f(\vec{a}\,)$ is a global maximum value.

A global maximum can occur at a local maximum or at a point on the boundary of the domain, denoted ∂D. Hence, to locate an absolute maximum or an absolute minimum, locate the local ones in the interior of D as above. Then parameterize the boundary of D and you will have a max-min problem of the sort you didi last year. Solve it and check which of your points has the largest (or smallest) value.

Returning to our example $f(x,y) = 2xy - 2x^3y - 2xy^3$ let us find the absolute maximum value if the domain is the square $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$.

We determined the local maxima and local minima above, so turn to the boundary. The boundary can be parameterized as four lines. Rather than repeat a similar calculation four times, observe all lines look like $x = at + b$, $y = ct + d$ so we need to study the functions $g(t) = f(at + b, ct + d)$ for various a, b, c and d.

Look at the line from $(0,0)$ to $(1,0)$: $x = t$, $y = 0$, $0 \leqslant t \leqslant 1$. Therefore $g(t) = 0$ on the entire line.

From $(1,0)$ to $(1,1)$: $x = 1$, $y = t$, $0 \leqslant t \leqslant 1$. Therefore $g(t) = 2t - 2t - 2t^3 = -2t^3$ so $g(t)$ has a min. at $t = 1$ and a max. at $t = 0$.

From $(1,1)$ to $(0,1)$, $x = t$, $y = 1$, $0 \leqslant t \leqslant 1$. Therefore $g(t) = 2t - 2t^3 - 2t = -2t^3$ so $g(t)$ has a min. at $t = 1$ and a max. at $t = 0$.

From $(0,1)$ to $(0,0)$, $x = 0$, $y = t$, $0 \leqslant t \leqslant 1$. Therefore $g(t) = 0$.

Absolute max. can occur at $(\frac{1}{2}, \frac{1}{2})$ and $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$, or anywhere on either line $(0,0) \leftrightarrow (1,0)$ or $(0,0) \leftrightarrow (0,1)$, where $f(x,y) = 0$, or at the end points of the other two lines where the maximum value is also 0. Hence the absolute maximum value is $\frac{1}{4}$ and it occurs only at the point $(\frac{1}{2}, \frac{1}{2})$.

There are no local minima in the square so the absolute minima must occur along the boundary. Along the two lines $(0,0) \leftrightarrow (1,0)$ and $(0,0) \leftrightarrow (0,1)$ the value of f is 0. Along the two lines $(1,0) \leftrightarrow (1,1)$ and $(0,1) \leftrightarrow (1,1)$ the minimum value is -2 and it occurs at $(1,1)$. Hence the absolute minimum value is -2 and it occurs only at $(1,1)$.

To astound your friends and confound your enemies you can repackage this discussion as:

On the square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$

$$-2 \leqslant 2xy - 2x^3y - 2xy^3 \leqslant \frac{1}{4}$$