

The Hessian and optimization

Let us start with two dimensions:

Let $f(x, y)$ be a function of two variables. We write the Taylor expansion around (x_0, y_0) . Write the vector $\vec{h} = \langle x - x_0, y - y_0 \rangle$. We will consider \vec{h} as a column vector when matrices are involved.

$$f(x, y) = f(x_0, y_0) + \underbrace{\frac{\partial f}{\partial x} \Big|_{(x_0, y_0)} (x - x_0) + \frac{\partial f}{\partial y} \Big|_{(x_0, y_0)} (y - y_0)}_{\nabla f \Big|_{(x_0, y_0)} \cdot \vec{h}} + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x^2} \Big|_{(x_0, y_0)} (x - x_0)^2 + \frac{\partial^2 f}{\partial x \partial y} \Big|_{(x_0, y_0)} (x - x_0)(y - y_0) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} \Big|_{(x_0, y_0)} (y - y_0)^2 + \dots}_{\frac{1}{2} \vec{h}^t H_f \Big|_{(x_0, y_0)} \vec{h}}$$

Here the H_f is the so called Hessian matrix

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

And as \vec{h} is a column vector, recall that \vec{h}^t is the transpose of \vec{h} , that is, the corresponding row vector, and so the computation becomes

$$\vec{h}^t H_f \vec{h} = \begin{bmatrix} (x - x_0) & (y - y_0) \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \begin{bmatrix} (x - x_0) \\ (y - y_0) \end{bmatrix}.$$

The trick is to understand what do these second derivatives tell us about the function. The best place to figure out what they mean is at a critical point, that is a point where $\nabla f = 0$. For simplicity suppose that $x_0 = y_0 = 0$, and $f(x_0, y_0) = 0$, and of course $\nabla f|_{(0,0)} = 0$. Then we have that

$$f(x, y) = \frac{1}{2} \vec{h}^t H_f \Big|_{(x_0, y_0)} \vec{h} + \dots,$$

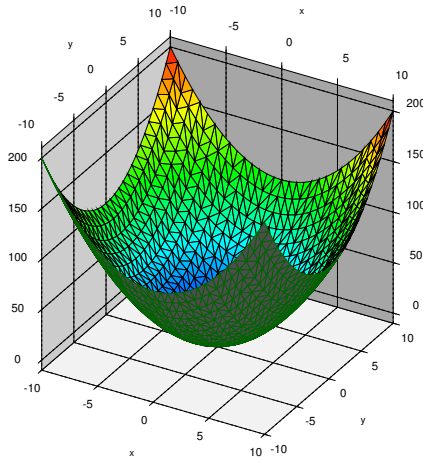
and so f behaves like a quadratic form near by the origin. Suppose that H_f is nondegenerate, that is, suppose that it has no zero eigenvalues. In that case it really behaves like one of three cases (after a linear change of coordinates).

Two positive eigenvalues:

The model case is when the matrix looks like

$$\frac{1}{2}H_f = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(we write half the matrix for simplicity). The function then becomes $x^2 + y^2$, and then the graph of the form looks like

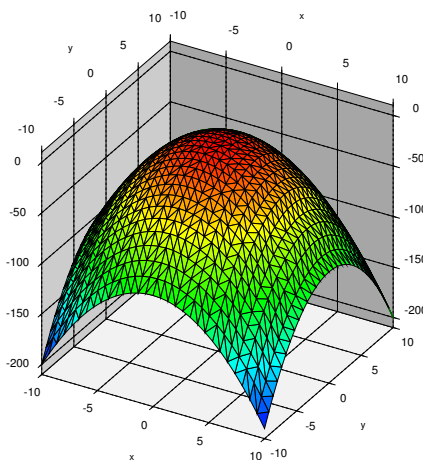


Two negative eigenvalues:

The model case is when the matrix looks like

$$\frac{1}{2}H_f = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The function then becomes $-x^2 - y^2$, and then the graph of the form looks like

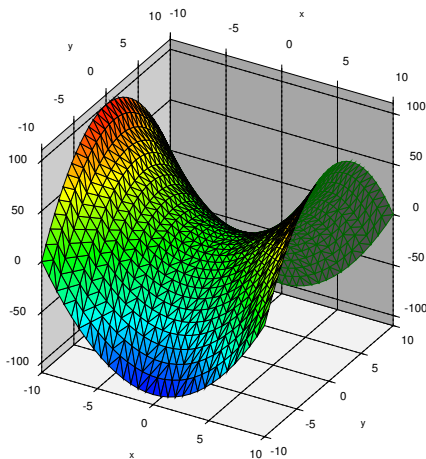


One positive and one negative eigenvalue:

The model case is when the matrix looks like

$$\frac{1}{2}H_f = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The function then becomes $x^2 - y^2$, and then the graph of the form looks like



Therefore, we if we wish to figure out if a critical point is a minimum or a maximum, (or a saddle), we look at the Hessian. If it is nondegenerate, we figure out what sign are the eigenvalues.

Computing eigenvalues is difficult, but there is a simple way to tell the signs. Notice

$$\det \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = 1, \quad \det \left(\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \right) = 1, \quad \det \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right) = -1.$$

The determinant of a matrix is a product of the eigenvalues. Therefore, if eigenvalues are of opposite signs, determinant of a 2×2 matrix is negative. Be careful, the reasoning is different in 3×3 , or larger, matrices.

It turns out that if you have two eigenvalues of the same sign, you can decide on the sign of the eigenvalues of a 2×2 matrix by just checking the upper left (or lower right) number in the matrix. If it is positive, then the matrix had two positive eigenvalues. One way you can see that is that you notice that if you above set $y = 0$, then $f(x, 0)$ is a function of x alone, and its second derivative is the upper

left number of the matrix. If it is positive, then the graph of the function must “curve up”, that is, it must be a minimum. Therefore, there must be at least one positive eigenvalue.

Let’s see an example. Take

$$f(x, y) = 3x^2 + 7xy + 5y^2.$$

The origin is a critical point, we just need to decide, is it a min/max or a saddle. The Hessian is

$$H_f = \begin{bmatrix} 6 & 7 \\ 7 & 10 \end{bmatrix}, \quad \text{and so} \quad \det(H_f) = \det \left(\begin{bmatrix} 6 & 7 \\ 7 & 10 \end{bmatrix} \right) = 6 \times 10 - 7 \times 7 = 11.$$

So the critical point at the origin is either a local minimum or a maximum. Since the upper left component is 6. It must be that the surface curves upwards, both eigenvalues are positive and the point is a minimum for f . In fact the eigenvalues are approximately 0.72 and 15.28, but notice that we didn’t have to check that.

Now suppose

$$g(x, y) = 3x^2 + 8xy + 5y^2.$$

Again the origin is a critical point. The Hessian is

$$H_g = \begin{bmatrix} 6 & 8 \\ 8 & 10 \end{bmatrix}, \quad \text{so} \quad \det(H_g) = \det \left(\begin{bmatrix} 6 & 8 \\ 8 & 10 \end{bmatrix} \right) = 6 \times 10 - 8 \times 8 = -4.$$

So the eigenvalues must be of opposite signs, and the surface must be a saddle point. In fact they are approximately -0.25 and 16.25 . Neither min nor max.

Degenerate Hessian

Consider

$$f(x, y) = x^2 + y^4.$$

It is not difficult to convince yourself that this function is always positive except at the origin, when it is zero. So the origin is a minimum.

However if you compute the Hessian you find

$$H_f = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{so} \quad \det(H_f) = 0.$$

The matrix is degenerate (it must have a zero eigenvalue since the determinant is zero). So we can't quite tell from the second derivative.

To see why we can't tell, try

$$g(x, y) = x^2 - y^4.$$

The graph of this is a saddle (Just notice that $g(x, 0) = x^2$ and $g(0, y) = -y^4$). If you compute the Hessian you find

$$H_g = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{so} \quad \det(H_g) = 0.$$

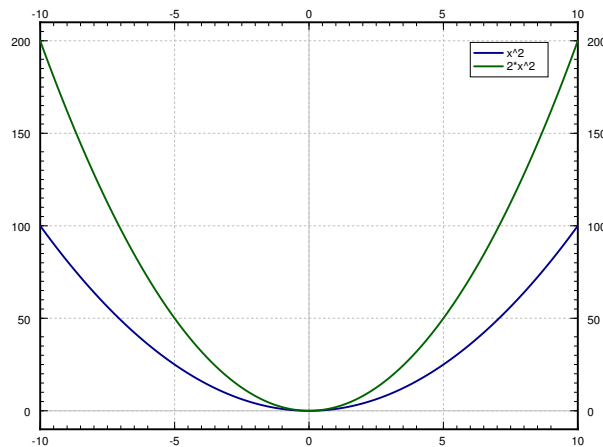
The Hessian is the same as for f above, so clearly we cannot read much from H_g and H_f . One thing that we can read off, is that the critical point is definitely not a maximum. That would require two negative eigenvalues, but one of the eigenvalues is in fact 2, which is positive (the other is the zero).

Peano's example

One has to be careful with trying to just check for extrema (min or max) on lines. Consider

$$f(x, y) = (y - x^2)(2x^2 - y).$$

This function, is always zero on the sets where $y = x^2$ and $y = 2x^2$. In fact, it is negative outside the space between these two lines and positive between the two lines.



If we restrict the function to any line, for example if we pick any angle θ , and consider the line through the origin at the angle θ to the positive real axis. That is,

the line given by $\gamma(t) = (t \cos(\theta), t \sin(\theta))$, then

$$f(\gamma(t)) = f(t \cos(\theta), t \sin(\theta)).$$

For any θ , $f(\gamma(t))$ has a strict maximum at $t = 0$. By a strict maximum we mean that the function is actually strictly less than zero for t near zero. In fact except for when $\theta = 0$ or $\theta = \pi$, we find

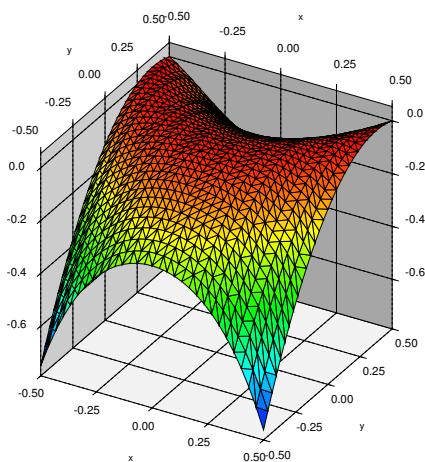
$$\frac{d^2}{dt^2} [f(\gamma(t))] < 0$$

so it is a maximum (a strict one). When $\theta = 0$, this corresponds to $y = 0$, and so

$$f(\gamma(t)) = -2t^4,$$

which clearly has a strict maximum.

But f is a saddle! It is just a “bent saddle.”



Clearly the function cannot have a min or a max since $f(0, 0) = 0$, but you also have points where f is positive or where f is negative arbitrarily near the origin.

A computation shows that

$$H_f = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}, \quad \text{so} \quad \det(H_f) = 0.$$

So we wouldn't be any smarter from looking at the Hessian.

The moral of the story is that if the Hessian has zero eigenvalues, you might have to get tricky to figure out what kind of critical point you have. There is no simple n th derivative test as there is for one variable functions.

More variables

Let's do this in 3 variables, with n variables being exactly the same idea. If $f(x, y, z)$ is a function of three variables, then the Hessian is the matrix

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

Again, a critical point is a minimum if H_f has 3 positive eigenvalues, it is a maximum if it has 3 negative eigenvalues and it is a saddle point if it has at least one positive and one negative eigenvalue.

Example: Let

$$f(x, y, z) = 3x^2 + xz + 2zy - z^2.$$

The Hessian matrix is

$$H_f = \begin{bmatrix} 6 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & -2 \end{bmatrix} \quad \text{so} \quad \det(H_f) = -24 < 0.$$

The determinant (the product of eigenvalues) is negative, so at least one eigenvalue is negative. So this cannot possibly be a minimum. But we cannot yet decide if it is a max or a saddle. However, we know we will be able to tell since it is nondegenerate. We just have to try harder. Either running through the linear algebra that you know or plugging into a computer or a calculator, you find that the eigenvalues are approximately $-3.31, 6.13, 1.18$. So one negative and two positive. The critical point must be a saddle.

Another example: Let

$$f(x, y, z) = -9x^2 + 6xy - 2y^2 - 2xz - 2z^2$$

The Hessian matrix is

$$H_f = \begin{bmatrix} -18 & 6 & 0 \\ 6 & -4 & -2 \\ 0 & -2 & -4 \end{bmatrix} \quad \text{so} \quad \det(H_f) = -72 < 0.$$

Again, negative, so at least one (or possibly all three) eigenvalues are negative. In this case one finds that the eigenvalues are approximately $-20.25, -0.70, -5.05$. So all three are negative.

There are other ways of figuring this out. For example, if the top left entry is negative, the first 2×2 principal submatrix has positive determinant, and the full matrix has negative determinant, then the eigenvalues are all negative. In this case

$$-18 < 0, \quad \det \left(\begin{bmatrix} -18 & 6 \\ 6 & -4 \end{bmatrix} \right) = 36 > 0, \quad \det \left(\begin{bmatrix} -18 & 6 & 0 \\ 6 & -4 & -2 \\ 0 & -2 & -4 \end{bmatrix} \right) = -72 < 0,$$

Basically you have to do some linear algebra to figure out the signs of the eigenvalues.

One common way to do this for 3 by 3 (and this generalizes to n by n) is the one we just used above. Given a symmetric matrix

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix},$$

it has three positive eigenvalues if

$$a > 0, \quad \det \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \right) > 0, \quad \text{and} \quad \det \left(\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \right) > 0.$$

And it has three negative eigenvalues if

$$a < 0, \quad \det \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \right) > 0, \quad \text{and} \quad \det \left(\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \right) < 0.$$

Computing determinants is a lot easier and a lot faster than computing the actual eigenvalues.