General Maxima and Minima, Quadratic Forms

Yosun Chang

September 21, 2005

The most general function of two variables can be expanded in the form of a Taylor series. Near the critical point of either a maxima, minima, or saddle-point, one can drop off the higher order terms in the expansion. One could move the origin arbitrarily, so that the critical point is renamed $f(0,0)$. One has, thus:

$$f(x,y) = f(0,0) + \frac{1}{2}f_{xx}x^2 + \frac{1}{2}f_{yy}y^2 + \frac{1}{2}f_{xy}xy + \frac{1}{2}f_{yx}yx + \mathcal{O}(3). \qquad (1)$$

$f(x,y)$ is the value of the function at any point (x,y). Its deviation from the critical point, call it Δf, is $\Delta f = f(x,y) - f(0,0)$.

The deviation from the critical point can be re-expressed in matrix form,

$$\Delta f = \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad (2)$$

which is also known as a *quadratic form*. (Marion and Thornton. Mechanics. Chapter 7)

Because $x^2, y^2 \geq 0$, if the cross-terms f_{xy} aren't there, one would be able to tell whether the deviation Δf is overall positive or negative when both $f_x x$ and $f_y y$ are of the same sign.

Because the coefficients f_{ij} can be expressed in the form of a matrix, as shown, one can transform the whole matrix quantity into principle coordinates, via the usual method (see the Review in the attached Appendix).

Recast in the basis of its principle coordinate system, the matrix has only diagonal entries, viz., its eigenvalues. Thus,

$$\Delta f = \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\left(f_{xx}x^2 + f_{yy}y^2\right). \qquad (3)$$

Thus, the following signs in the coefficients f_{ii} determines everything:

$$\boxed{f_{xx}, f_{yy} < 0 \Rightarrow f(x,y) < f(0,0) \Rightarrow f(0,0) \text{ is a local max}}$$

$$\boxed{f_{xx}, f_{yy} > 0 \Rightarrow f(x,y) > f(0,0) \Rightarrow f(0,0) \text{ is a local min}}$$

1 Example

Usually, the critical point isn't $(0, 0)$ (however, since we can move our "origin" anywhere, so the above is just as general). You can find the critical point by taking the first derivative of your function and setting the resulting expression equal to 0.

$f = x^3 - 2x + y^2$

$f_x = 3x^2 - 2 = 0 \Rightarrow x = \pm\sqrt{2/3}$ and $f_y = 2y = 0 \Rightarrow y = 0$

Now, the f_{ij} matrix (also known as the Hessian matrix) looks like this:

$$f_{ij} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 6x & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} \pm 6\sqrt{2/3} & 0 \\ 0 & 2 \end{pmatrix} \tag{4}$$

It just happens that the function I randomly conjured happens to already be in its principle coordinate system, and thus we immediately have our eigenvalues of $6x$ and 0. The last step above comes from plugging in x, as determined by setting $f_x = 0$.

So, this means that at the point $(\sqrt{3/2}, 0)$, we have a local minimum (since both $f_{ij} > 0$). At the point $(-\sqrt{32}, 0)$, we have a max.

But, if the function *isn't already* in its principle coordinate system, then you'd have to find its eigenvalues and pop them into the diagonals of your final f_{ij} matrix. Details below.

2 Review: Transforming to Principle Coordinates

The matrix A may be any funky square matrix filled with ugly entries. In the basis of its principle coordinate system, however, it will have only diagonal elements, viz., its eigenvalues.

2.1 Finding Eigenvalues

$|A - \lambda| = 0$. Solve for λ. A $n \times n$ matrix has n λ's.

2.2 Finding Eigenvectors

Plug in the lambda's into the original equation: $A|x\rangle = \lambda|x\rangle$. x is a simple linear equation. Pick numbers that obey the equation, and you get your eigenvector corresponding to λ.

3 References

Shankar. Basic Training in Mathematics. p 274ff