1 Connections To Calculus

Quadratic forms come up in calculus when we look at the second derivatives of functions. Remember in single variable calculus that every differentiable function $f : \mathbb{R} \to \mathbb{R}$ has a Taylor expansion at the point p:

$$f(p) + \frac{f'(p)}{1!}(x-p) + \frac{f''(p)}{2!}(x-p)^2 + \cdots$$

By truncating this sum we get a polynomial that is supposed to approximate f near p. The more of the terms of the infinite series we keep in our polynomial, the better the approximation. The first order approximation, $L_1(x) = f(p) + f'(p)(x-p)$ just encodes the slope of the function at p. Its graph is the tangent line at $(p, f(p))$. The first order is good, but we might want to look at the second order term to learn more about f. We do this most often when $f'(p) = 0$ (such a p is called a critical pt). At critical points by examining the sign of the second derivative we were able to say whether f had a max or min. Remember that if the second derivative was zero we couldn't say anything.

OK, the story is much the same for differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$, except now the Taylor expansion gets messier. It needs to take into account all the possible ways to take partial derivatives. Here is the beginning of the Taylor expansion at $\vec{p} = \begin{bmatrix} p_x \\ p_y \end{bmatrix}$ in the case of $\mathbb{R}^2$:

$$f(\vec{p}) + \frac{\partial f}{\partial x}(\vec{p})(x-p_x) + \frac{\partial f}{\partial y}(\vec{p})(y-p_y) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\vec{p})(x-p_x)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y \partial x}(\vec{p})(x-p_x)(y-p_y)$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x \partial y}(\vec{p})(y-p_y)(x-p_x) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(\vec{p})(y-p_y)^2 + \cdots$$

which we can rewrite succintly as

$$f(\vec{p}) + D_{\vec{p}}f\left(\begin{bmatrix} x \\ y \end{bmatrix} - \vec{p}\right) + \frac{1}{2}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \vec{p}\right)^T \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}_{\vec{p}} \left(\begin{bmatrix} x \\ y \end{bmatrix} - \vec{p}\right)$$

Up 'til now we have only been interested in the first order approximation given by $D_{\vec{p}}f$. But since we're growing up...fast, I might add...it is time to examine the second order term especially when the first order term vanishes (i.e. $D_{\vec{p}}f = [0 \cdots 0]$).

Studying that beasty looking second order term

The first thing to mention is that we call that nasty looking matrix containg all the second order partial derivatives the *Hessian* matrix, and for short we write $Hess(f)_{\vec{p}}$ for the Hessian matrix evaluated at $\vec{p}$. The next thing to notice is that since partial derivatives commute, the matrix

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}_{\vec{p}}$$

is symmetric. Well, now we're in business since we've become little experts in symmetric matrices and their associated quadratic forms. I should just add that

$$\left(\begin{bmatrix} x \\ y \end{bmatrix} - \vec{p}\right)^T Hess(f)_{\vec{p}} \left(\begin{bmatrix} x \\ y \end{bmatrix} - \vec{p}\right)$$

is the quadratic form

$$\begin{bmatrix} x \\ y \end{bmatrix}^T Hess(f)_{\vec{p}} \begin{bmatrix} x \\ y \end{bmatrix}$$

translated over to the point $\vec{p}$.

Alrighty, if this matrix has all positive eigenvalues, then we know its graph is a concave up parabaloid. Provided $D_{\vec{p}}f$ vanishes, the second order approximation to f is a parabaloid whose minimum value is $f(p)$ attained at the point $\vec{p}$. We then conclude that f has a local minimum at $\vec{p}$. Similarly if this matrix has all negative eigenvalues, then we know its graph is a concave down parabaloid. Provided $D_{\vec{p}}f$ vanishes, the second order approximation to f has a local maximum at $\vec{p}$. If the Hessian has both negative and positive eigenvalues and $D_{\vec{p}}f$ vanishes, then f is neither a local min or max, but locally looks like a saddle point.

IMPORTANT: if the Hessian has 0 as an eigenvalue, then we cannot give a qualitative description of f near $\vec{p}$ just by looking at the Hessian.

I know this may seem complicated, but the story actually parallels the story for the single variable case. When you are getting confused, try to think about the single variable case and figure out things from there.

Example 1 Let

$$f : \mathbb{R}^2 \to \mathbb{R}$$
$$(x, y) \mapsto x^2 + y^2 - 8x + 6y + 2$$

Find the critical points of f and describe f near these points.

solution 1) The critical points are those points $\vec{p}$ where

$$[0\ 0] = D_{\vec{p}}f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}_{\vec{p}}$$
$$= \begin{bmatrix} 2x - 8 & 2y + 6 \end{bmatrix}_{\vec{p}}.$$

So we must solve for all $\vec{p} = \begin{bmatrix} x \\ y \end{bmatrix}$ that satisfy the equations

$$0 = 2x - 8$$
$$0 = 2y + 6.$$

Clearly the only solution is (4,-3).

2) OK, to analyze what f is like near (4,-3) we'll want the second order approximation to f. This requires knowing the Hessian, whose entries are the second order partials of f:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}(2x - 8) = 2$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}(2y+6) = 2$$

and the partials. Since partials commute/the Hessian is symmetric, we only have to bother computing one of them.

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}(2x-8) = 2$$

From this we conclude that the Hessian at (4,-3) is

$$= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}_{(4,-3)} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Since the eigenvalues of this Hessian are clearly all positive, we know that f must have a local minimum at (4,-3).

The second order approximation to f at (4,-3) is

$$f((4,-3)) + \frac{1}{2}\begin{bmatrix} x-4 \\ y+3 \end{bmatrix}^T \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x-4 \\ y+3 \end{bmatrix}$$

$$= -23 + \frac{1}{2}\begin{bmatrix} x-4 & y+3 \end{bmatrix} \begin{bmatrix} 2(x-4) \\ 2(y+3) \end{bmatrix} = -23 + (x-4)^2 + (y+3)^2$$

Alright...another one. Same shit, just a slightly more complicated function f. **Example 2** This time let

$$f : \mathbb{R}^2 \to \mathbb{R}$$

$$(x,y) \mapsto \frac{xy}{x^2+1} - \frac{2}{5}y$$

1) Again, the critical points are those points $\vec{p}$ where

$$[0\ 0] = D_{\vec{p}} f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}_{\vec{p}}$$

Using the product rule on f thinking of it as $(x^2+1)^{-1} \cdot xy - \frac{2}{5}y$.

$$\frac{\partial f}{\partial x} = 2x \cdot -1 \cdot \frac{1}{x^2+1^2} \cdot xy + \frac{y}{x^2+1} = \frac{-x^2+1}{(x^2+1)^2} y$$

$$\frac{\partial f}{\partial y} = \frac{x}{x^2+1} - \frac{2}{5}$$

Plugging these into $\begin{bmatrix} 0 & 0 \end{bmatrix} = D_{\vec{p}} f$ yields the equations

$$0 = \frac{-x^2+1}{(x^2+1)^2} y$$

$$0 = \frac{x}{x^2+1} - \frac{2}{5}$$

The first equation involves both x and y and looks a little scary. The second one is more manageable. Adding $\frac{2}{5}$ to both sides and then multiplying both sides by x^2+1 (which is never zero btw) yields

$$\frac{2}{5}x^2 + \frac{2}{5} = x$$

This is a quadratic equation which hopefully we know how to solve. I got x=2 or $\frac{1}{2}$. OK, now we plug these values of x into the first equation we either

$$0 = \frac{-3}{25}y \text{ and } 0 = \frac{12}{25}y \text{ respectively}$$

In both cases, we find that y=0. So we conclude that f has critical points at (2,0) and ($\frac{1}{2}$,0).

2) Lets analyze what f is like near (2,0) first. Again, we'll need the second order approximation to f. This requires knowing the Hessian, whose entries are the second order partials of f:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{-x^2+1}{(x^2+1)^2}y\right)$$

and I'm actually not going to bother figuring it out, because I see that whatever I get, it will have a factor of y. Since we'll be evaluating these second order partials at places where y=0, I see that I'll get zero.

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(2y+6) = 2$$

and the partials. Since partials commute/the Hessian is symmetric, we only have to bother computing one of them. Lets pick the easier of the two given the partial derivatives we've already computed:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{-x^2+1}{(x^2+1)^2}y\right) = \frac{-x^2+1}{(x^2+1)^2}$$

From this we conclude that the Hessian at (2,0) is

$$= \begin{bmatrix} 0 & \frac{-x^2+1}{(x^2+1)^2} \\ \frac{-x^2+1}{(x^2+1)^2} & 0 \end{bmatrix}_{(2,0)} = \begin{bmatrix} 0 & -\frac{3}{25} \\ -\frac{3}{25} & 0 \end{bmatrix}.$$

Lets solve for the eigenvalues.

$$0 = \begin{vmatrix} -\lambda & -\frac{3}{25} \\ -\frac{3}{25} & -\lambda \end{vmatrix} = \lambda^2 - \left(\frac{3}{25}\right)^2$$

So the eigenvalues are $\pm\frac{3}{25}$. Since at least one eigenvalue is positive and one negative, we conclude that near (2,0) f has a saddle. It has neither a local max nor min at (2,0). The second order polynomial that approximates f near (2,0) is:

$$f((2,0)) + \frac{1}{2}\begin{bmatrix} x-2 \\ y \end{bmatrix}^T \begin{bmatrix} 0 & -\frac{3}{25} \\ -\frac{3}{25} & 0 \end{bmatrix} \begin{bmatrix} x-2 \\ y \end{bmatrix}$$

$$= 0 + \frac{1}{2}\begin{bmatrix} x-2 & y \end{bmatrix} \begin{bmatrix} -\frac{3}{25}y \\ -\frac{3}{25}(x-2) \end{bmatrix} = -\frac{3}{25}(x-2)y$$

Whew, almost done. Now we do the same for f at the point ($\frac{1}{2}$,0). We've already computed the second order partials, we just need to evaluate at ($\frac{1}{2}$,0).

$$Hess_{(\frac{1}{2},0)}(f) = \begin{bmatrix} 0 & \frac{12}{25} \\ \frac{12}{25} & 0 \end{bmatrix}$$

Similarly to before, we find the eigenvalues are $\pm\frac{12}{25}$, so we conclude that near ($\frac{1}{2}$,0) f looks like a saddle point. It has neither a local max nor min at ($\frac{1}{2}$,0).

The second order polynomial that approximates f near ($\frac{1}{2}$,0) is:

$$f((\frac{1}{2},0)) + \frac{1}{2}\begin{bmatrix} x-\frac{1}{2} \\ y \end{bmatrix}^T \begin{bmatrix} 0 & \frac{12}{25} \\ \frac{12}{25} & 0 \end{bmatrix} \begin{bmatrix} x-\frac{1}{2} \\ y \end{bmatrix}$$

$$= 0 + \frac{1}{2}\begin{bmatrix} x-2 & y \end{bmatrix} \begin{bmatrix} \frac{12}{25}y \\ \frac{12}{25}(x-\frac{1}{2}) \end{bmatrix} = \frac{12}{25}(x-\frac{1}{2})y$$