7Z. Applications

(see also chapter V from oldnotes.pdf)

One reason for studying more advanced topics in mathematics is to obtain more insight into things that are already known. We shall consider two such applications: Describing surfaces in $\mathbb{R}^3$ defined by second degree polynomials, and justifying the Second Derivative Test for relative absolute minima in multivariable calculus.

Quadric hypersurfaces in $\mathbb{R}^n$

These are defined by equations of the form

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{k=1}^{n} b_k x_k + c = 0$$

where not all of the a_{ij}'s are zero, and we might as well assume $a_{ij} = a_{ji}$ (we can always modify the coefficients to get such an equations; in general, if $A_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$ then $A_{ij} = A_{ji}$ and $\sum A_{ij} x_i x_j + \sum b_k x_k + c = 0$ is equivalent to the original equation.).

We can rewrite the equation in matrix form with $\underset{n\times n}{A=(a_{ij})}$ symmetric, $\underset{1\times n}{B=(b_k)}$, $\underset{n\times 1}{X=(x_k)}$:

$$ {}^TX A X + B X + c = 0 . $$

What happens if we make a linear change of variables? Express it as $X=PY$ where P is invertible (so $Y=QX$, where $Q=P^{-1}$). The result is

$$ {}^TY({}^TPAP)Y + (\overbrace{BP}^{B'})Y + c = 0 $$

Now choose P orthogonal such that the columns of P are eigenvectors for A. If $P=(\overset{\text{COLUMNS}}{p_1 \ldots p_n})$ then $AP=(Ap_1 \ldots Ap_n)=(\lambda_1 p_1 \ldots \lambda_n p_n)=$ PD where D is the diagonal matrix $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

Since $AP=PD$ and P is orthogonal, we have ${}^TPAP={}^TPPD=D$.

Therefore we may make a linear change of variables to rewrite the defining equation in the form

$$\sum_j d_j y_j^2 + \sum_j b_k' y_j + c = 0 .$$

Further simplification.

Let r = nonzero d_j's, and assume that the y's are ordered so that the nonzero d's precede the zero d's. We can then complete squares $z_j = y_j + \frac{b_j}{2d_j}$ if $j \leq r$, $z_j = y_j$ if $j > r$,

to obtain the following equivalent equation:

$$\sum_{j=1}^{r} d_j z_j^2 + \sum_{k=r+1}^{n} b_k' z_k + c' = 0.$$

maybe some $b_k' = 0$

One can take this even further, but this is simple enough for now.

Examples when $n=3$

$$ax^2+by^2+cz^2-k=0 \qquad a,b,c\neq 0$$

The exact shape depends upon the signs of a, b, c, k

$$ax^2+by^2+cz-k=0 \qquad a,b\neq 0$$

Same comment as above (with the possibility that $c=0$ now).

$$ax^2+by+cz-k=0 \qquad a\neq 0$$

Similarly, but now maybe $b=0$ too. In fact, we can change variables so that $c=0$ here. (No proof).

Second derivative test for critical points

U = open region in $\mathbb{R}^2$, $f : U \to \mathbb{R}$
has continuous second partial derivatives.
Local maxima + minima must have $\nabla f(p) = 0$.
$\frac{\partial f}{\partial x_i}(p) = 0$ all i.

Look at 2nd order partial derivatives to recognize relative maxima, minima or neither. Generalize to n dimensions now.

Multivariable Taylor approximation

$a \in U$ + $|h| < r \Rightarrow a + h \in U$. Then

$$f(a+h) = f(a) + \nabla f(a) \cdot h + {}^{T}h \, Hf(a + t^* h) \, h$$

where $0 \leq t^* \leq 1$ and

$Hf(y)$ is the symmetric matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(y) \right)$.

Hessian

Now specialize to the case $n = 2$.

For functions of one variable, if $f'(a)=0$ then we can tell if a is a relative maximum or minimum if $f''(a) \neq 0$; if $f''(a) > 0$ one has a relative minimum, if $f''(a) < 0$ one has a relative maximum. There are analogous results in two variables.

Relative minima $\nabla f(a) = 0$ and

$$\frac{\partial^2 f}{\partial x^2}(a) > 0 \qquad \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x \partial y}(a) \\ \frac{\partial^2 f}{\partial x \partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{vmatrix} > 0$$

$\Rightarrow$ f has a relative minimum at a.

Proof. The Taylor approximation is

$$f(a+h) = f(a) + {}^{T}h \cdot Hf(a+t^*h) \cdot h$$

This will be a relative minimum if the eigenvalues of the Hessian are positive. By previous lectures, this happens if $\frac{\partial^2 f}{\partial x^2}(a+t^*h) > 0$, $\begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}(a+t^*h) > 0$.

Now $\frac{\partial^2 f}{\partial x^2}$ and $\begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}$ are continuous, so if $\|k\| <$ some small r^*, these are positive for $(a+k)$.

Therefore the Taylor approximation implies

$$f(a) < f(a+h) \text{ if } h \neq 0.$$

Relative maxima $\nabla f(a) = 0$ and

$$\frac{\partial^2 f}{\partial x^2}(a) < 0 \qquad \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}(a) > 0 \implies$$

f has a relative maximum at a.

Sketch of proof. f has a relative maximum $\iff$ $-f$ has a relative minimum, which happens if

$$-\frac{\partial^2 f}{\partial x^2}(a) > 0 \qquad \begin{vmatrix} -\frac{\partial^2 f}{\partial x^2} & -\frac{\partial^2 f}{\partial x \partial y} \\ -\frac{\partial^2 f}{\partial x \partial y} & -\frac{\partial^2 f}{\partial y^2} \end{vmatrix}(a) > 0$$

equivalent to

$$\frac{\partial^2 f}{\partial x^2}(a) < 0$$

this determinant is equal to

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}(a) \, (> 0).$$

$$\left(\text{since } \begin{vmatrix} -a & -b \\ -c & -d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \right)$$

Saddle point unstable — relative maximum in one direction, relative minimum in another.

Typical example $f(x,y) = y^2 - x^2$ (saddle surface)

$(0,0)$ is a relative maximum in the direction of the x-axis, and it is a relative minimum in the direction of the y-axis.

~~Proof~~ Suppose $\nabla f(a) = 0$ and

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} (a) < 0.$$ Then f has a saddle point at a.

Proof. Since the determinant of Hf at a is negative, it follows that the eigenvalues of $Hf(a)$ have opposite signs. Let u_+ + u_- be orthonormal eigenvectors with positive and negative eigenvalues λ_+, λ_-, and define functions

$$\alpha(t) = f(a + tu_+)$$
$$\beta(t) = f(a + tu_-)$$
for $|t|$ small.

The chain rule then implies that

$\alpha'(0) = \beta'(0) = 0$, and further computations

with the chain rule yield

$$\alpha''(0) = 2\lambda_+ > 0 > 2\lambda_- = \beta''(0).$$

Therefore α has a relative minimum and β has a relative maximum at zero.

If the determinant of $Hf(a)$ is zero, then no conclusion can be drawn.