MATH 223. Quadratic Forms, Conic Sections. Richard Anstee

When faced with a quadratic function such as \( x^2 + 3xy + y^2 + 2yz + 2z^2 \) we discover that we can write it using a matrix:

\[
x^2 + 3xy + y^2 + 2yz + 2z^2 = [x \ y \ z] \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} [x \ y \ z]
\]

and then we make the interesting observation that we can do this with a symmetric matrix:

\[
x^2 + 3xy + y^2 + 2yz + 2z^2 = [x \ y \ z] \begin{bmatrix} 1 & 3/2 & 0 \\ 3/2 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} [x \ y \ z]
\]

The symmetric matrix makes this so much easier to analyze. Sometimes this takes a non diagonalizable matrix to a symmetric matrix.

\[
x^2 + 4xy + y^2 = [x \ y] \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} [x \ y] = [x \ y] \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} [x]
\]

Let \( A \) be a symmetric matrix with \( AM = MD \) for an orthogonal matrix \( M \) (with \( M^T = M^{-1} \)) and a diagonal matrix \( D \). In this case

\[
\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}
\]

We apply this with \( x^T Ax = x^T MDM^T x = z^T Dz \) where \( z = M^T x = [u \ v]^T \) (or \( x = Mz \)). Then

\[
x^T Ax = x^T M^T DMx = u^T Du \quad \text{for} \quad u = Mx.
\]

A change of variable allows us to do diagonalization in this setting. This question does have \( M^T = M \) which is unusual. Using the change of variables \([u \ v]^T = M^T x\) or more explicitly

\[
u = \frac{1}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y \quad \text{and} \quad v = \frac{1}{\sqrt{2}} x - \frac{1}{\sqrt{2}} y,
\]

then we have our original expression \( x^2 + 4xy + y^2 = 3u^2 - v^2 \), where perhaps the second expression in \( u, v \) is simpler.

An exam question was to determine sketch the family of curves given by

\[
x^2 + 8xy - 5y^2 = t
\]

for various \( t \). These curves would be called Conic Sections and would arise from the intersection of a plane with a double cone (e.g. \( \{(x, y, z) : x^2 + y^2 = |z|\} \)).

Local Extrema

Given a function \( f(x_1, x_2, \ldots, x_n) \) of \( n \) variables, one would look for critical points. For example \( 0 \) is critical when

\[
\frac{\partial f}{\partial x_1} \big|_{x=0} = 0, \quad \frac{\partial f}{\partial x_2} \big|_{x=0} = 0, \quad \cdots, \quad \frac{\partial f}{\partial x_n} \big|_{x=0} = 0
\]

Let

\[
A = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_1 x_3} & \cdots \\ \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 x_3} & \cdots \\ \frac{\partial^2 f}{\partial x_1 x_3} & \frac{\partial^2 f}{\partial x_2 x_3} & \frac{\partial^2 f}{\partial x_3^2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}
\]
We have that $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ so the matrix is symmetric. But it is also true that the partial derivatives provide the coefficients for the second degree Taylor polynomial (centred at $x = 0$) in the $n$ variables. We have that

$$\frac{\partial^2}{\partial x_i \partial x_j} x_i x_j = 1 \text{ while } \frac{\partial^2}{\partial x_i \partial x_i} x_i^2 = 2.$$

We then compute that $f \approx \frac{1}{2} x^T A x + f(0)$ (using our hypothesis that the first derivatives are 0 at $x = 0$).

Now $x^T A x = x M D M^T x = (M^T x)^T D (M^T x)$. Let $z = M^T x$. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues so that these are the diagonal entries of $D$, then

$$x^T A x = z^T D z = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2.$$

The point $x = 0$ is a *local minimum* if $x^T A x > 0$ for $x \neq 0$. This is true if and only if all the eigenvalues of $A$ are positive! Similarly the point $x = 0$ is a *local maximum* if $x^T A x < 0$ and so if all the eigenvalues of $A$ are negative. The orthogonality of the eigenspaces is used to provide the appropriate change of variables $z = M^T x$.

Interestingly there are quick ways to test if all the eigenvalues are positive that do not require computing the eigenvalues due to Sylvester.

There are special ways to test for these properties for symmetric matrices (Sylvester’s Law of Inertia) which I won’t prove here.