MATH 223. Quadratic Forms, Conic Sections. Richard Anstee

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

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

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

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

The symmetric matrix makes this so much easier to analyze. In this particular case you can compute (I used Wolfram Alpha) that eigenvalues are $\lambda_1 = 1$, $\lambda_2 = (1/2)(2 + \sqrt{13})$, $\lambda_3 = (1/2)(2 - \sqrt{13})$. Eigenvectors are $\mathbf{v}_1 = (-2/3, 0, 1)^T$, $\mathbf{v}_2 = (3/2, \sqrt{13}/2, 1)^T$, $\mathbf{v}_3 = (3/2, -\sqrt{13}/2, 1)^T$. The eigenvectors are orthogonal, not yet orthonormal. They make the axes for the picture.

We readily deduce that for $f(x, y, x) = const > 0$, that we have cross sections in v_1, v_2 plane being eliptical with the ellipse growing as the coordinate in v_3 direction is more distant from 0. Some call this a hyperboloid of one sheet. For $f(x, y, x) = const < 0$, there will be no points with the coordinate in v_3 being 0 but as you move away from origin it will again be an ellipses growing in size as the coordinate in v_3 direction is more distant from the origin. Sometimes this takes a non diagonalizable matrix to a symmetric matrix.

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

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 $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T MDM^T\mathbf{x} = \mathbf{z}^T D\mathbf{z}$ where $\mathbf{z} = M^T\mathbf{x} = [u\, v]^T$ (or $\mathbf{x} = M\mathbf{z}$). Then

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T MDM^T\mathbf{x} = \mathbf{u}^T D\mathbf{u} \quad \text{for } \mathbf{u} = M^T\mathbf{x}.$$

A change of variable allows us to do diagonalization in this setting. This question does have $M^T = M$ which follows from our theorem on the orthogonal diagonalzation of symmetric matrices.. Using the change of variables $[u\, v]^T = M^T\mathbf{x}$ or more explicitly $u = \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y$ and $v = \frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y$, 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.

We could view these transformation as a general completing the square idea. Admittedly we need a transformation to remove linear terms but that is comparitivel straighforward after the diagonalization (or perhaps before).

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

$$x^2 + 8xy - 5y^2 = t$$

for various t. As t varies you get nice hyperbolas except that for $t = 0$ you get two lines (the axes of the hyperbola for the cases with $t \neq 0$. These curves would be called *Conic Sections* and would arise from the intesection of a plane with a double cone (e.g. $\{(x, y, z) : x^2 + y^2 = |z|\}$).

In these curve problems, the orthonormal matrices have the virtue of not disturbing the shape of the curve but merely rotating and perhaps reflecting. General change of basis matriceswill often disturb such curves greatly . Consider the white and blue coordinates in our handout. An example would be to imagine the circle $x^2 + y^2 = 1$ in white coordinates and try to see it as an ellipse in blue coordinates and similarly consider the circle $v_1^2 + v_2^2 = 1$ in blue coordinates and try to see it as an ellipse in white coordinates.

You do need to have a little familiarity with these curves such as

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

which is the ellipse with semi axes a (along x-axis) and b (along y-axis). Of course using plotting software would help visualize in 2 and 3 dimensions.

Local Extrema

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

$$\frac{\partial}{\partial x_1} f(\mathbf{x})|_{\mathbf{x}=\mathbf{0}} = 0, \quad \frac{\partial}{\partial x_2} f(\mathbf{x})|_{\mathbf{x}=\mathbf{0}} = 0, \quad \cdots, \frac{\partial}{\partial x_n} f(\mathbf{x})|_{\mathbf{x}=\mathbf{0}} = 0$$

Let

$$A = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} & \cdots \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \cdots \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3 \partial x_3} & \cdots \\ & \vdots & & \end{bmatrix}$$

This is called the Hessian. 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 $\mathbf{x} = \mathbf{0}$) 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 $f \approx \frac{1}{2}\mathbf{x}^T A\mathbf{x} + f(\mathbf{0})$ (using our hypothesis that the first derivatives are 0 at $\mathbf{x} = \mathbf{0}$).

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

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

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

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