II. Diagonalisation of real quadratic forms

Let V be a n-dimensional vector space over $\mathbb{R}$ and let $v_1, \ldots, v_n$ be a basis for V.

Definition. A *quadratic form* $q : V \to \mathbb{R}$ is a function of the form

$$q(v) = \mathbf{x}^t A \mathbf{x}$$

where A is a symmetric matrix and $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ where $v = x_1 v_1 + \cdots + x_n v_n$.

Examples. (1) If

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

then

$$\begin{aligned} q(x, y) &= q(xv_1 + yv_2) \\ &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= ax^2 + 2bxy + cy^2. \end{aligned}$$

(2) If

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix},$$

then

$$\begin{aligned} q(x, y, z) &= q(xv_1 + yv_2 + zv_3) \\ &= \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz. \end{aligned}$$

Remark. In fact any function $q : V \to \mathbb{R}$ of the form

$$q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} b_{ij} x_i x_j$$

is a quadratic form. The corresponding symmetric matrix is then $A = (a_{ij})$ where $a_{ii} = b_{ii}$ and $a_{ij} = a_{ji} = \frac{1}{2} b_{ij}$ if $i < j$.

Example. The function $q(x, y, z) = 2x^2 - y^2 + 3xz + 4yz$ is the quadratic form corresponding to the matrix

$$A = \begin{pmatrix} 2 & 0 & 3/2 \\ 0 & -1 & 2 \\ 3/2 & 2 & 0 \end{pmatrix}.$$

Remark. These crop up frequently in mathematics. For example in Analysis when one is looking for local extreme points. For $f : \mathbb{R}^n \to \mathbb{R}$ let p be some critical point (i.e. where $\nabla f = 0$). The Hessian

$$H_f(p) = \left(\frac{\delta^2 f}{\delta x_i \, \delta x_j} \right)_p$$

is then a quadratic form.

It is clearly of interest to try to choose the basis $v_1, \dots, v_n$ so that the quadratic form has the simplest form. So let us next think of the effect the change of the basis has. If $w_1, \dots, w_n$ is another basis for V and P is the base change matrix that takes from $w_1, \dots, w_n$ to $v_1, \dots, v_n$. Then

$$v = y_1 w_1 + \dots + y_n w_n = x_1 v_1 + \dots + x_n v_n$$

if $\mathbf{x} = P\mathbf{y}$. So

$$q(v) = (P\mathbf{y})^t AP\mathbf{y} = \mathbf{y}^t (P^t AP)\mathbf{y}.$$

We say that two matrices A and B are *congruent* if $B = P^t AP$ for some invertible real matrix P. Notice that this is not the same as being similar (where we need $B = P^{-1}AP$). The first one is an equivalence relation we use when we deal with quadratic forms and the latter we use when dealing with linear operators.

We are now going to see that for a given symmetric matrix A, we can always find a congruent matrix B that is diagonal (and in fact has only $1, -1$ and 0 on the diagonal).

To see this we apply the Spectral Theorem. Choose a inner product which makes $v_1, \dots, v_n$ an orthonormal basis (that is we define the inner product such that $\langle v_i, v_j \rangle = \delta_{ij}$ where δ_{ij} is the Kronecker symbol). Let $\alpha : V \to V$ be the self-adjoint operator that is represented by the symmetric matrix A. Then

$$q(v) = \langle v, \alpha(v) \rangle.$$

By the Spectral Theorem, we can choose a new orthonormal basis $w_1, \dots, w_n$ of eigenvectors of α so that now

$$q(v) = \mathbf{x}^t B\mathbf{x}$$

where $\mathbf{x}$ gives us the coordinates of v in the new basis and if P is the base change matrix from the new basis to the old then

$$B = P^{-1}AP = P^t AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

(Notice that B is both congruent and similar to A). Next we rescale the new basis by letting

$$Q = \begin{pmatrix} q_1 & & \\ & \ddots & \\ & & q_n \end{pmatrix}$$

where

$$q_i = \begin{cases} 1/\sqrt{|\lambda_i|}, & \text{if } \lambda_i \neq 0 \\ 1 \ \text{ if } \lambda_i = 0 \end{cases}$$

Then $C = Q^t BQ$ is a diagonal matrix with only $-1, 1$ and 0 on the diagonal. Reordering the rescaled basis we obtain.

Theorem 7.9 *(Sylvesters law of inertia) Every symmetric real matrix A is congruent to some*

$$\begin{pmatrix} I & & \\ & -I & \\ & & 0 \end{pmatrix}.$$

One can show further that the size r of I and the size s of $-I$ are invariants of the quadratic form.

Terminology. The number $r + s$ is called the *rank* of q and $r - s$ is called the *signature* of q.

One says that q is *positive definite* if $r = n$ and $s = 0$ and *negative definite* if $r = 0$ and $s = -n$.

Remark. For example if $H_f(p)$ is positive definite one has local min at p and if $H_f(p)$ is negative definite then one has local max at p.