Remark on quadratic forms

Minhyong Kim

November 26, 2008

I wish to mention an important point of view about quadratic forms that wasn't quite what was presented in the lectures or the notes.

I believe the origin of the techniques we are discussing is the study of *quadratic equations*, say

$$f(x, y) = x^2 + 4xy + 5y^2 = c$$

or

$$q(x, y, z) = 3x^2 + 6y^2 + 11z^2 + 8xy + 16yz + 10xz = c$$

It is in fact rather surprising that linear algebra has something to say about this, since the subject, after all, is concerned in the main with the study of *linear equations*. The possibility of this extended application has to do with the procedure (which must have seemed quite mysterious to early practitioners) of associating a matrix to the quadratic form. That is,

$$f(x, y) = (x, y)A(x, y)^t$$

for the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

(Notice here that I am writing out the row vector (x, y) and putting the corresponding column vector as $(x, y)^t$, in contrast to the usual convention in class. This is for ease of typing.) and

$$q(x, y, z) = (x, y, z)B(x, y, z)^t$$

for

$$B = \begin{pmatrix} 3 & 4 & 5 \\ 4 & 6 & 8 \\ 5 & 6 & 11 \end{pmatrix}$$

At this point, it might be reasonable to review how to get at the matrices starting from the quadratic forms. For $f(x, y)$ it is quite easy once we are told we should search for an A, since we know that A is of the form

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$$

A simple substitution will reveal $\alpha = 1, \gamma = 5, \beta = 4/2 = 2$. (Of course the insight is exactly to start the search.) For q, the substitution may get a bit tedious, so it might be more useful to recall that the (i, j) entry is given by the value of the corresponding corresponding bilinear form on the pair of vectors (e_i, e_j) (here the e_i refer to the standard basis) which is

$$(1/2)(q(e_i + e_j) - q(e_i) - q(e_j))$$

The diagonal entries are therefore quickly computed as

$$q(1, 0, 0) = 3, q(0, 1, 0) = 6, q(0, 0, 1) = 11$$

that is, the coefficients of x^2, y^2, and z^2. And then, we compute the $(1,2)$ entry as

$$(q(1,1,0) - q(1,0,0) - q(0,1,0))/2 = (17 - 3 - 6)/2 = 4,$$

the $(2,3)$ entry is

$$(q(0,1,1) - q(0,1,0) - q(0,0,1))/2 = (33 - 6 - 11)/2 = 8$$

while the $(1,3)$ entry is

$$(q(1,0,1) - q(1,0,0) - q(0,0,1))/2 = (24 - 3 - 11)/2 = 5$$

This is all we need to compute because the entries below the diagonal can be filled in using the symmetry of the matrix. I hope it's convincing that even in any number of variables

$$g(x_1, x_2, \ldots, x_n) = \Sigma_{i<j} g_{ij} x_i x_j$$

it is straightforward to compute a symmetric $n \times n$ matrix G such that

$$g(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n) G (x_1, x_2, \ldots, x_n)^t$$

This week's coursework requires you to compute a matrix with respect a different (i.e., not the standard one) basis $B = \{b_1, \ldots, b_n\}$, in which case the entries will be just

$$[g(b_i + b_j) - g(b_i) - g(b_j)]/2$$

How does all this help with quadratic equations?

The reason is all very classical and concerns the process of *changing variables.* For $f(x,y)$ for example, completing the square gives us

$$f(x,y) = (x + 2y)^2 + y^2$$

so that introducing the variables $x' = (x + 2y)$, $y' = y$ puts the polynomial into simplified form

$$f(x', y') = (x')^2 + (y')^2$$

Thus, if you are interested in plane curves, this will tell you that the geometry of the curve defined by

$$x^2 + 4xy + 5y^2 = 1$$

is 'essentially the same' as the geometry of the circle. For quadratic polynomials in two variables, completing the square is clearly an elementary technique for simplifying the structure of equations.

However, in three variables already, let alone an arbitrary number, it is not a priori obvious how to carry out a similar simplification, or if it is even possible. Instead, we perform some double operations on the matrix B:

$$B = \begin{pmatrix} 3 & 4 & 5 \\ 4 & 6 & 8 \\ 5 & 8 & 11 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 5 & 8 & 11 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 & 5 \\ 1 & 1 & 3 \\ 5 & 3 & 11 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 1 & 3 \\ 3 & 1 & 5 \\ 5 & 3 & 11 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 3 \\ 1 & 3 & 5 \\ 3 & 5 & 11 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 2 \\ 3 & 5 & 11 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \\ 3 & 2 & 11 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

If we keep track of the elementary matrices used for the column operation portion of each step, we see that they are

$$E_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_3 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_4 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_5 = \begin{pmatrix} 1 & 0 & -0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Since the diagonal matrix at the end is

$$D = E_5^t(E_4^t(E_3^t(E_2^t(E_1^t B E_1)E_2)E_3)E_4)E_5$$

we see that the transition matrix is

$$P = E_1E_2E_3E_4E_5 = \begin{pmatrix} -1 & 2 & 1 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

(We could have arrived at P at the same time as the diagonalization by just performing the successive column operations starting with the identity matrix.) That is,

$$D = P^tBP,$$

and

$$B = (P^{-1})^tDP^{-1}.$$

The inverse matrix can be found in the standard way, say using row operations, to yield

$$P^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore,

$$q(x, y, z) = (x, y, z)B(x, y, z)^t = (x, y, z)(P^{-1})^t DP^{-1}(x, y, z)^t = [P^{-1}(x, y, z)^t]^t D[P^{-1}(x, y, z)^t].$$

Meanwhile

$$P^{-1}(x, y, z)^t = (x + 2y + 3z, x + y + z, z)^t$$

So we see that introducing the variables

$$x' = x + 2y + 3z, y' = x + y + z, z' = z$$

will allow us to express the quadratic form as

$$q(x', y', z') = (x')^2 + 2(y')^2$$

To briefly consider again the geometry, we see that the geometry of the solutions to

$$3x^2 + 6y^2 + 11z^2 + 8xy + 16yz + 10xz = 1,$$

which may look quite complicated, is essentially the same as the elliptic cylinder

$$(x')^2 + 2(y')^2 = 1$$

Thus, a problem that can look complicated from the view of classical substitution was resolved readily using simple matrix operations. That is, the double operations, relying crucially on the matrix representation, give a routine procedure for finding the correct substitution, as opposed to elaborate computations with variables. Recall in this regard that the process of Gaussian elimination for solving linear systems also asks us to briefly forget the variables and work purely with the coefficient matrix.

Our substitutions above occurred over the rational numbers $\mathbb{Q}$, and hence, could have been used to describe rational solutions to the equations. Some of you will meet this fact again in a course on Diophantine equations. However, the use of square roots and the the attendant canonical forms tells us that over the reals, the geometry of any equation of the form

$$\Sigma_{i<j} g_{ij} x_i x_j = 1$$

can be reduced to

$$x_1^2 + x_2^2 + \cdots x_r^2 - x_{r+1}^2 - x_{r+2}^2 - \cdots + x_{r+s}^2 = 1$$

while over the complex numbers, it can even be reduced to

$$x_1^2 + x_2^2 + \cdots x_r^2 + x_{r+1}^2 + x_{r+2}^2 + \cdots + x_{r+s}^2 = 1$$

You may find this fact useful if you end up studying some *algebraic geometry*, a powerful subject with numerous applications to number theory, information theory, robotics, and theoretical physics.

Incidentally, if you consider the slightly meta-mathematical question of why mere linear algebra could have been used for non-linear equations, the answer lies properly in the subject of *multi-linear algebra* or the study of *tensors*. A more elementary answer is that we discussed just a single quadratic polynomial rather than a whole system.