MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 16

In the first half of this module, matrices were used to represent linear maps. We shall temporarily take the outlook which views symmetric matrices in the spirit of one of the proofs from last week, and looks at the associated bilinear forms (and quadratic forms). Let us start with a motivating example.

Motivation

Example 1. Consider a function $f(x_1,\ldots,x_n)$ of n scalar real arguments. Let $\mathbf{x}^0=x_1^0e_1+\cdots+x_n^0e_n$ be a point in $\mathbb{R}^n$, and assume that the function f is smooth enough to consider its Taylor series to order two near the point $\mathbf{x}^0$:

$$f(\mathbf{x})=f(\mathbf{x}^0)+\sum_{i=1}^{k}(x_i-x_i^0)\frac{\partial f}{\partial x_i}(\mathbf{x}^0)+\frac{1}{2}\sum_{i,j=1}^{n}(x_i-x_i^0)(x_j-x_j^0)\frac{\partial^2 f}{\partial x_i x_j}(\mathbf{x}^0)+o(|\mathbf{x}-\mathbf{x}^0|^2).$$

Suppose that we would like to know whether f attains its locally minimal/maximal value at $\mathbf{x}^0$. Then, since in the first order of magnitude we have

$$f(\mathbf{x})=f(\mathbf{x}^0)+\sum_{i=1}^{k}(x_i-x_i^0)\frac{\partial f}{\partial x_i}(\mathbf{x}^0)+o(|\mathbf{x}-\mathbf{x}^0|),$$

we conclude that a necessary condition is $\frac{\partial f}{\partial x_i}(\mathbf{x}^0)=0$ for all i, that is the *gradient* of f vanishes at $\mathbf{x}^0$. In this case, we have

$$f(\mathbf{x})=f(\mathbf{x}^0)+\frac{1}{2}\sum_{i,j=1}^{n}(x_i-x_i^0)(x_j-x_j^0)\frac{\partial^2 f}{\partial x_i x_j}(\mathbf{x}^0)+o(|\mathbf{x}-\mathbf{x}^0|^2),$$

so the difference between $f(\mathbf{x})$ and $f(\mathbf{x}^0)$, when $\mathbf{x}$ is close to $\mathbf{x}^0$, is approximately equal to $\frac{1}{2}q(\mathbf{x}-\mathbf{x}^0)$, where

$$q(\mathbf{y})=a_{11}y_1^2+2a_{12}y_1y_2+\cdots+2a_{1n}y_1y_n+y_2^2+2a_{23}y_2y_3+\cdots+a_{nn}y_n^2,$$

where for brevity we denote $a_{ij}=\frac{\partial^2 f}{\partial x_i x_j}(\mathbf{x}^0)$; we have $a_{ij}=a_{ji}$ whenever the function f is smooth enough. The function $q(\mathbf{y})$ is a very typical example of a *quadratic form*.

Bilinear and quadratic forms

Definition 1. Let V be a vector space. A function $q\colon V\to\mathbb{R}$ is said to be a *quadratic form* if for some basis $e_1,\ldots,e_n$ of V we have

$$q(x_1e_1+\cdots+x_ne_n)=\sum_{1\le i,j\le n}a_{ij}x_ix_j,$$

where a_{ij} are some scalars. In other words, the value of q on a vector is a quadratic polynomial in coordinates of a vector.

Remark 1. It is easy to see that if the condition from the definition holds for some basis, then it holds for any basis, since coordinates relative to different bases are related by transition matrices in a linear way.

One simple example of a quadratic form is

$$q(x) = x_1^2 + \cdots + x_n^2.$$

In general, if V is a Euclidean vector space then $q(x) = (x, x)$ is certainly a quadratic form. This can be generalised: every *bilinear form* gives rise to a quadratic form.

Definition 2. Let V be a vector space. A function $V \times V \to \mathbb{R}$, $v_1, v_2 \mapsto b(v_1, v_2)$ is called a bilinear form if for all vectors v, v_1, v_2 the following conditions are satisfied:

$$b(c_1 v_1 + c_2 v_2, v) = c_1 b(v_1, v) + c_2 b(v_2, v) \quad \text{and} \quad b(v, c_1 v_1 + c_2 v_2) = c_1 b(v, v_1) + c_2 b(v, v_2).$$

A bilinear form is said to be *symmetric* if $b(v_1, v_2) = b(v_2, v_1)$ for all v_1, v_2. A symmetric bilinear form is said to be *positive definite*, if $b(v, v) \geqslant 0$ for all v, and $b(v, v) = 0$ only for $v = 0$.

In these words, a function of two vector arguments is a scalar product if and only if it is bilinear, symmetric, and positive definite.

Another important class of bilinear forms is given by skew-symmetric ones: a bilinear form is said to be *skew-symmetric* if $b(v_1, v_2) = -b(v_2, v_1)$ for all v_1, v_2. Those appear very frequently in differential geometry and advanced classical mechanics.

Remark 2. Generalising what we proved about scalar products, for every bilinear form b and every basis $e_1, \ldots, e_n$ of V, we have

$$b(x_1 e_1 + \ldots + x_n e_n, y_1 e_1 + \ldots + y_n e_n) = \sum_{i,j=1}^{n} b_{ij} x_i y_j,$$

where $b_{ij} = b(e_i, e_j)$. Moreover, writing this sum as

$$\sum_{i=1}^{n} x_i \sum_{j=1}^{n} b_{ij} y_j,$$

we see that

$$b(x_1 e_1 + \ldots + x_n e_n, y_1 e_1 + \ldots + y_n e_n) = x^T B y,$$

where the $n \times n$-matrix B has entries b_{ij}. (Strictly speaking, $x^T B y$ is a 1×1-matrix, but that is essentially the same as a single number.)

Every bilinear form b gives rise to a quadratic form by putting $q(x) = b(x, x)$, for example, the bilinear form

$$b(x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = 2 x_1 y_2$$

gives rise to a quadratic form $2 x_1 x_2$, and the bilinear form

$$b(x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_2 + x_2 y_1$$

gives rise to the same quadratic form. It turns out that the reconstruction of b from q is unique if we assume that b is symmetric; in this case the reconstruction formula is

$$b(v, w) := \frac{1}{2}(q(v + w) - q(v) - q(w)).$$

One celebrated example of a quadratic form is $q(x_1, x_2, x_3, t) = x_1^2 + x_2^2 + x_3^2 - t^2$ on the Minkowski space $\mathbb{R}^4$, it is used in special relativity theory. This (sort of) motivates the next few results. We shall now formulate them (in order to use them in the next homework), and next week we shall discuss their proofs in detail.

Four theorems about signed sums of squares

Theorem 1 (Canonical form)**.** *Let q be a quadratic form on a vector space V. There exists a basis $e_1, \ldots, e_n$ of V for which the quadratic form q becomes a signed sum of squares:*

$$q(x_1 e_1 + \cdots + x_n e_n) = \sum_{i=1}^{n} \varepsilon_i x_i^2,$$

where all numbers ε_i are either 1 or -1 or 0.

Theorem 2 (Law of inertia)**.** *In the previous theorem, the triple (n_+, n_-, n_0), where $n_\pm$ is the number of ε_i equal to ± 1, and n_0 is the number of ε_i equal to 0, does not depend on the choice of the basis $e_1, \ldots, e_n$. This triple is often referred to as the* signature *of the quadratic form q.*

Let $B = (b_{ij})$ be the matrix of a given symmetric bilinear form b on V. We shall now discuss some methods of computing the signature of b via the matrix elements of B. We denote by B_k the $k \times k$-matrix whose entries are b_{ij} with $1 \leqslant i, j \leqslant k$, that is the top left corner submatrix of B, and put $\Delta_0 = 1$ and $\Delta_k := \det(B_k)$ for $1 \leqslant k \leqslant n$.

Theorem 3 (Jacobi theorem)**.** *Suppose that for all $i = 1, \ldots, n$ we have $\Delta_i \neq 0$. Then the number of 1's, 0's and -1's in the canonical form is equal to the number of positive numbers, zeros, and negative numbers among the numbers $\frac{\Delta_{k-1}}{\Delta_k}$, $k = 1, \ldots, n$.*

Theorem 4 (Sylvester theorem)**.** *The given symmetric bilinear form is positive definite if and only if*

$$\Delta_k > 0 \quad \textit{for all} \quad k = 1, \ldots, n.$$