Math 395. Quadratic spaces over $\mathbf{R}$

1. Algebraic preliminaries

Let V be a vector space over a field F. Recall that a *quadratic form* on V is a map $Q : V \to F$ such that $Q(cv) = c^2 Q(v)$ for all $v \in V$ and $c \in F$, and such that the symmetric pairing $\beta_Q : V \times V \to F$ defined by $\beta_Q(v, w) = Q(v + w) - Q(v) - Q(w)$ is *bilinear*. (The explicit coordinatized description, sometimes presented as the definition, will be given shortly.) A *quadratic space* over F is a pair (V, Q) consisting of a vector space V over F and a quadratic form Q on V.

Note that $\beta_Q(v, v) = Q(2v) - 2Q(v) = 2Q(v)$, so as long as $1 + 1 \neq 0$ in F we can run the procedure in reverse: for any symmetric bilinear pairing $B : V \times V \to F$, $Q_B(v) = B(v, v)$ is a quadratic form on V and the two operations $Q \mapsto B_Q = \beta_Q/2$ and $B \mapsto Q_B$ are inverse bijections between quadratic forms on V and symmetric bilinear forms on V. Over general fields, one cannot recover Q from β_Q. (Example: $q(x) = x^2$ and $Q(x) = 0$ on $V = F$ have $\beta_q = 0 = \beta_Q$ when $2 = 0$ in F, yet $q \neq 0$.) When $2 \neq 0$ in F, we say that Q is *non-degenerate* exactly when the associated symmetric bilinear pairing $B_Q = \beta_Q/2 : V \times V \to F$ is perfect (that is, the associated self-dual linear map $V \to V^\vee$ defined by $v \mapsto B_Q(v, \cdot) = B_Q(\cdot, v)$ is an isomorphism, or more concretely the "matrix" of B_Q with respect to a basis of V is invertible). In other cases (with $2 \neq 0$ in F) we say Q is *degenerate*.

If $\dim V = n$ is finite and positive, and we choose a basis $\{e_1, \dots, e_n\}$ of V, then for $v = \sum x_i e_i$ we have

$$Q(v) = Q(\sum_{i<n} x_i e_i + x_n e_n) = Q(\sum_{i<n} x_i e_i) + Q(x_n e_n) + \beta_Q(\sum_{i<n} x_i e_i, x_n e_n),$$

and bilinearity gives the last term as $\sum_{i<n} c_{in} x_i x_n$ with $c_{in} = \beta_Q(e_i, e_n) \in F$. Also, $Q(x_n e_n) = c_{nn} x_n^2$ with $c_{nn} = Q(e_n) \in F$. Hence, inducting on the number of terms in the sum readily gives

$$Q(\sum x_i e_i) = \sum_{i \leq j} c_{ij} x_i x_j$$

with $c_{ij} \in F$, and conversely any such formula is readily checked to define a quadratic form. Note also that the c_{ij}'s are uniquely determined by Q (and the choice of basis): the formula forces $Q(e_i) = c_{ii}$, and then setting $x_i = x_j = 1$ for some $i < j$ and setting all other $x_k = 0$ gives $Q(e_i + e_j) = c_{ij} + c_{ii} + c_{jj}$, so indeed c_{ij} is uniquely determined. One could therefore say that a quadratic form "is" a homogeneous quadratic polynomial in the linear coordinates x_i's, but this coordinatization tends to hide underlying structure and make things seem more complicated than necessary, much like in the study of "matrix algebra" without the benefit of the theory of vector spaces and linear maps.

Example 1.1. Suppose $2 \neq 0$ in F, so we have seen that there is a bijective correspondence between symmetric bilinear forms on V and quadratic forms on V; this bijection is even linear with respect to the evident linear structures on the sets of symmetric bilinear forms on V and quadratic forms on V (using pointwise operations; $(a_1 B_1 + a_2 B_2)(v, v') = a_1 B_1(v, v') + a_2 B_2(v, v')$, which one checks is symmetric bilinear, and $(a_1 Q_1 + a_2 Q_2)(v) = a_1 Q_1(v) + a_2 Q_2(v)$ which as a function from V to F is checked to be a quadratic form). Let us make this bijection concrete, as follows. In class we saw that if we fix an ordered basis $\mathbf{e} = \{e_1, \dots, e_n\}$ of V then we can describe a symmetric bilinear $B : V \times V \to F$ in terms of the matrix $[B] = {}_{\mathbf{e}^\vee}[\varphi_\ell]_{\mathbf{e}} = (b_{ij})$ for the "left/right-pairing" map $\varphi_\ell = \varphi_r$ from V to $V^\vee$ defined by $v \mapsto B(v, \cdot) = B(\cdot, v)$, namely $b_{ij} = B(e_j, e_i) = B(e_i, e_j)$. However, in terms of the dual linear coordinates $\{x_i = e_i^*\}$ we have just seen that we can uniquely write $Q_B : V \to F$ as $Q_B(v) = \sum_{i \leq j} c_{ij} x_i(v) x_j(v)$. What is the relationship between the c_{ij}'s and the b_{ij}'s?

We simply compute: for $v = \sum x_i e_i$, bilinearity of B implies that $Q_B(v) = B(v,v)$ is given by

$$\sum x_i x_j B(e_i, e_j) = \sum_i B(e_i, e_i) x_i^2 + \sum_{i<j} (B(e_i, e_j) + B(e_j, e_i)) x_i x_j = \sum_i b_{ii} x_i^2 + \sum_{i<j} 2 b_{ij} x_i x_j,$$

where $b_{ij} = B(e_j, e_i) = B(e_i, e_j) = b_{ji}$. Hence, $c_{ii} = b_{ii}$ but for $i < j$ we have $c_{ij} = 2b_{ij} = b_{ij} + b_{ji}$. Thus, for B and Q that correspond to each other, given the polynomial $[Q]$ for Q with respect to a choice of basis of V, we "read off' the symmetric matrix $[B]$ describing B (in the same linear coordinate system) as follows: the ii-diagonal entry of $[B]$ is the coefficient of the square term x_i^2 in Q, and the "off-diagonal" matrix entry b_{ij} for $i \neq j$ is given by *half* the coefficient for $x_i x_j = x_j x_i$ appearing in $[Q]$ (recall $2 \neq 0$ in F). For example, if $Q(x,y,z) = x^2 + 7y^2 - 3z^2 + 4xy + 3xz - 5yz$ then the corresponding symmetric bilinear form B is computed via the symmetric matrix

$$[B] = \begin{pmatrix} 1 & 2 & 3/2 \\ 2 & 7 & -5/2 \\ 3/2 & -5/2 & -3 \end{pmatrix}.$$

Going in the other direction, if someone hands us a *symmetric* matrix $[B] = (b_{ij})$ then we "add across the main diagonal" to compute that the corresponding homogeneous quadratic polynomial $[Q]$ is $\sum_i b_{ii} x_i^2 + \sum_{i<j} (b_{ij} + b_{ji}) x_i x_j = \sum_i b_{ii} x_i^2 + \sum_{i<j} 2b_{ij} x_i x_j$.

It is an elementary algebraic fact (to be proved in a moment) for any field F in which $2 \neq 0$ that, relative to some basis $\mathbf{e} = \{e_1, \dots, e_n\}$ of V, we can express Q in the form $Q = \sum \lambda_i x_i^2$ for some scalars $\lambda_1, \dots, \lambda_n$ (some of which may vanish). In other words, we can "diagonalize" Q, or rather the "matrix" of B_Q (and so the property that some λ_i vanishes is equivalent to the intrinsic property that Q is degenerate). To see why this is, we note that Q is uniquely determined by B_Q (as $1 + 1 \neq 0$ in F) and in terms of B_Q this says that the basis consists of vectors $\{e_1, \dots, e_n\}$ that are mutually perpendicular with respect to B_Q (i.e., $B_Q(e_i, e_j) = 0$ for all $i \neq j$). Thus, we can restate the assertion as the general claim that if $B : V \times V \to F$ is a symmetric bilinear pairing then there exists a basis $\{e_i\}$ of V such that $B(e_i, e_j) = 0$ for all $i \neq j$. To prove this we may induct on $\dim V$, the case $\dim V = 1$ being clear. In general, if $n = \dim V > 1$ then we first choose a nonzero $e_n \in V$ and we note that $v \mapsto B(e_n, v) = B(v, e_n)$ is a linear functional on V. Thus, either its kernel is a hyperplane or is all of V, and so either way the kernel contains a hyperplane H. We use induction for B restricted to $H \times H$ to find a suitable $e_1, \dots, e_{n-1}$ that, together with e_n, solve the problem.

2. Some generalities over $\mathbf{R}$

Now assume that $F = \mathbf{R}$. Since all positive elements of $\mathbf{R}$ are squares, after first passing to a basis of V that "diagonalizes" Q (which, as we have seen, is a purely algebraic fact), we can rescale the basis vectors using $e'_i = e_i / \sqrt{|\lambda_i|}$ when $\lambda_i \neq 0$ to get (upon reordering the basis)

$$Q = {x'}_1^2 + \dots + {x'}_r^2 - {x'}_{r+1}^2 - \dots - {x'}_{r+s}^2$$

for some $r, s \geq 0$ with $r + s \leq \dim V$. Let $t = \dim V - r - s \geq 0$ denote the number of "missing variables" in such a diagonalization (so $t = 0$ if and only Q is non-degenerate). The value of r here is just the number of λ_i's which were positive, s is the number of λ_i's which were negative, and t is the number of λ_i's which vanish. The values r, s, t a priori may seem to depend on the original choice of ordered basis $\{e_1, \dots, e_n\}$.

To shed some light on the situation, we introduce some terminology that is specific to the case of the field $\mathbf{R}$. The quadratic form Q is *positive-definite* if $Q(v) > 0$ for all $v \in V - \{0\}$, and Q is *negative-definite* if $Q(v) < 0$ for all $v \in V - \{0\}$. Since $Q(v) = B_Q(v,v)$ for all $v \in V$, clearly if Q is either positive-definite or negative-definite then Q is non-degenerate. In terms of the diagonalization with all coefficients equal to ± 1 or 0, positive-definiteness is equivalent to the condition $r = n$ (and so this possibility is coordinate-independent), and likewise negative-definiteness is equivalent to the condition $s = n$. In general we define the *null cone* to be

$$C = \{v \in V \,|\, Q(v) = 0\},$$

so for example if $V = \mathbf{R}^3$ and $Q(x, y, z) = x^2 + y^2 - z^2$ then the null cone consists of vectors $(x, y, \pm\sqrt{x^2 + y^2})$ and this is physically a cone (or really two cones with a common vertex at the origin and common central axis). In general C is stable under scaling and so if it is not the origin then it is a (generally infinite) union of lines through the origin; for $\mathbf{R}^2$ and $Q(x, y) = x^2 - y^2$ it is a union of two lines.

Any vector v not in the null cone satisfies exactly one of the two possibilities $Q(v) > 0$ or $Q(v) < 0$, and we correspondingly say (following Einstein) that v is *space-like* or *time-like* (with respect to Q). The set V^+ of space-like vectors is an open subset of V, as is the set V^- of time-like vectors. These open subsets are disjoint and cover the complement of the null cone. In the preceding example with $Q(x, y) = x^2 - y^2$ on $V = \mathbf{R}^2$, V^+ and V^- are each disconnected (as drawing a picture shows quite clearly). This is atypical:

Lemma 2.1. *The open set V^+ in V is non-empty and path-connected if $r > 1$, with r as above in terms of a diagonalizing basis for Q, and similarly for V^- if $s > 1$.*

Proof. By replacing Q with $-Q$ if necessary, we may focus on V^+. Obviously V^+ if non-empty if and only if $r > 0$, so we may now assume $r \geq 1$. We have

$$Q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2$$

with $r \geq 1$ and $0 \leq s \leq n - r$. Choose $v, v' \in V^+$, so $x_j(v) \neq 0$ for some $1 \leq j \leq r$. We may move along a line segment contained in V^+ to decrease all $x_j(v)$ to 0 for $j > r$ (check!), and similarly for v', so for the purposes of proving connectivity we can assume $x_j(v) = x_j(v') = 0$ for all $j > r$. If $r > 1$ then v and v' lie in the subspace $W = \mathrm{span}(e_1, \ldots, e_r)$ of dimension $r > 1$ on which Q has positive-definite restriction. Hence, $W - \{0\} \subseteq V^+$, and $W - \{0\}$ is path-connected since $\dim W > 1$. ■

The basis giving such a diagonal form is simply a basis consisting of r space-like vectors, s time-like vectors, and $n - (r + s)$ vectors on the null cone such that all n vectors are B_Q-perpencidular to each other. In general such a basis is rather non-unique, and even the subspaces

$$V_{+,\mathbf{e}} = \mathrm{span}(e_i \,|\, \lambda_i > 0), \quad V_{-,\mathbf{e}} = \mathrm{span}(e_i \,|\, \lambda_i < 0)$$

are *not* intrinsic. For example, if $V = \mathbf{R}^2$ and $Q(x, y) = x^2 - y^2$ then we can take $\{e_1, e_2\}$ to be either $\{(1, 0), (0, 1)\}$ or $\{(2, 1), (1, 2)\}$, and we thereby get different spanning lines. Remarkably, it turns out that the values

$$r_{\mathbf{e}} = |\{i \,|\, \lambda_i > 0\}| = \dim V_{+,\mathbf{e}}, \quad s_{\mathbf{e}} = |\{i \,|\, \lambda_i < 0\}| = \dim V_{-,\mathbf{e}}, \quad t_{\mathbf{e}} = |\{i \,|\, \lambda_i = 0\}| = \dim V - r_{\mathbf{e}} - s_{\mathbf{e}}$$

are independent of the choice of "diagonalizing basis" $\mathbf{e}$ for Q. One thing that is clear right away is that the subspace

$$V_{0,\mathbf{e}} = \mathrm{span}(e_i \,|\, \lambda_i = 0)$$

is actually intrinsic to V and Q: it is the set of $v \in V$ that are B_Q-perpendicular to the entirety of V: $B_Q(v, \cdot) = 0$ in $V^\vee$. (Beware that this is not the set of $v \in V$ such that $Q(v) = 0$; this latter set is the null cone C, and it is never a linear subspace of V when it contains nonzero points.)

3. Algebraic proof of well-definedness of the signature

Theorem 3.1. *Let V be a finite-dimensional $\mathbf{R}$ vector space, and Q a quadratic form on V. Let $\mathbf{e}$ be a diagonalizing basis for Q on V. The quantities $\dim V_{+,\mathbf{e}}$ and $dim V_{-,\mathbf{e}}$ are independent of $\mathbf{e}$.*

We'll prove this theorem using algebraic methods in a moment (and a longer, but more illuminating, proof by connectivity considerations will be given later in the course). In view of the intrinsic nature of the number of positive coefficients and negative coefficients in a diagonal form for Q (even though the specific basis giving rise to such a diagonal form is highly non-unique), we are motivated to make the:

Definition 3.2. Let Q be a quadratic form on a finite-dimensional $\mathbf{R}$-vector space V. We define the *signature* of (V, Q) (or of Q) to be the ordered pair of non-negative integers (r, s) where $r = \dim V_{+,\mathbf{e}}$ and $s = \dim V_{-,\mathbf{e}}$ respectively denote the number of positive and negative coefficients for a diagonal form of Q. In particular, $r + s \leq \dim V$ with equality if and only if Q is non-degenerate.

The signature is an invariant that is intrinsically attached to the finite-dimensional quadratic space (V, Q) over $\mathbf{R}$. In the study of quadratic spaces over $\mathbf{R}$ with the fixed dimension, it is really the "only" invariant. Indeed, we have:

Corollary 3.3. *Let (V, Q) and (V', Q') be finite-dimensional quadratic spaces over $\mathbf{R}$ with the same finite positive dimension. The signatures coincide if and only if the quadratic spaces are isomorphic; i.e., if and only if there exists a linear isomorphism $T : V \simeq V$ with $Q'(T(v)) = Q(v)$ for all $v \in V$.*

This corollary makes precise the fact that the signature and dimension are the only isomorphism class invariants in the *algebraic* classification of finite-dimensional quadratic spaces over $\mathbf{R}$. However, even when the signature is fixed, there is a lot more to do than mere algebraic classification. There's a lot of geometry in the study of quadratic spaces over $\mathbf{R}$, so the algebraic classification via the signature is not the end of the story. We now prove the corollary, granting Theorem 3.1, and then we will prove Theorem 3.1.

Proof. Assume such a T exists. If $\mathbf{e}$ is a digonalizing basis for Q, clearly $\{T(e_i)\}$ is a diagonalizing basis for Q' with the same diagonal coefficients, whence Q' has the same signature as Q. Conversely, if Q and Q' have the same signatures (r, s) there exist ordered bases $\mathbf{e}$ and $\mathbf{e}'$ of V and V' such that in terms of the corresponding linear coordinate systems $x_1, \dots, x_n$ and $x'_1, \dots, x'_n$ we have

$$Q = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2, \quad Q' = {x'}_1^2 + \cdots + {x'}_r^2 - {x'}_{r+1}^2 - \cdots - {x'}_{r+s}^2.$$

Note in particular that

$$Q(\sum a_i e_i) = \sum_{i=1}^{r} a_i^2 - \sum_{i=r+1}^{s} a_i^2 = Q'(\sum a_i e'_i)$$

for all i. Thus, if $T : V \to V'$ is the linear map determined by $T(e_i) = e'_i$ then T sends a basis to a basis. Thus, T is a linear isomorphism, and also

$$Q'(T(\sum a_i e_i)) = Q'(\sum a_i e'_i) = Q(\sum a_i e_i).$$

In other words, $Q' \circ T = Q$, as desired. ■

Now we turn to the proof of the main theorem stated above.

Proof. Let $V_0 = \{v \in V \mid B_Q(v, \cdot) = 0\}$. Let $\mathbf{e} = \{e_1, \dots\}$ be a diagonalizing basis of V for Q, with $Q = \sum \lambda_i x_i^2$ relative to $\mathbf{e}$-coordinates, where $\lambda_1, \dots, \lambda_{r_\mathbf{e}} > 0$, $\lambda_{r_\mathbf{e}+1}, \dots, \lambda_{r_\mathbf{e}+s_\mathbf{e}} < 0$, and $\lambda_i = 0$ for $i > r_\mathbf{e} + s_\mathbf{e}$. Clearly $V_0 = V_{0,\mathbf{e}}$.

Now we consider the subspaces $V_{+,\mathbf{e}}$ and $V_{-,\mathbf{e}}$. Since these subspaces (along with $V_0 = V_{0,\mathbf{e}}$) are given as the span of parts of the basis $\mathbf{e}$ (chopped up into three disjoint pieces, some of which may be empty), we have a decomposition

$$V = V_{+,\mathbf{e}} \oplus V_{-,\mathbf{e}} \oplus V_0$$

with $B_Q(v^+, v^-) = 0$ for all $v^+ \in V_{+,\mathbf{e}}$, $v^- \in V_{-,\mathbf{e}}$ (due to the diagonal shape of B_Q relative to the $\mathbf{e}$-coordinates). Also, it is clear from the diagonal form of Q that $Q(v^+) \geq 0$ for all $v^+ \in V_{+,\mathbf{e}}$ with $Q(v^+) = 0$ if and only if $v^+ = 0$ (since Q on the subspace $V_{+,\mathbf{e}}$ is presented as the sum of squares of basis coordinates). Likewise, we have $Q(v^-) \leq 0$ for $v^- \in V^-$, with equality if and only if $v^- = 0$.

In other words, the diagonalizing basis $\mathbf{e}$ for Q on V gives rise to a decomposition

$$V = V^+ \oplus V^- \oplus V_0$$

of V into subspaces, with

- Q positive definite on V^+;
- Q negative definite on V^-;
- $B_Q(v^+, v^-) = 0$ for all $v^+ \in V^+$, $v^- \in V^-$ (and V_0 is the intrinsic subspace discussed above).

Note that *conversely* if we are given any such direct sum decomposition of V with these listed properties, then upon choosing a diagonalizing basis for $Q|_{V^+}$ on V^+, for $Q|_{V^-}$ on V^-, and any basis of V_0, we obtain a *diagonalizing* basis for Q on V. The point here is that the combined basis on V yields a coordinatized version of Q not only without "cross-terms" involving pairs of linear coordinates coming from the same subspace (V^+, V^-, or V_0), but also without cross terms involving a linear coordinate from one of these subspaces and a linear coordinate from another one. This is exactly due to the fact that these three subspaces are "mutually perpendicular" with respect to B_Q (note that elements of V_0 are even "perpendicular" to everything in V). This point of view provides a more intrinsic basis-free description of exactly the problem we aim to study.

Namely, we assume we are given two direct sum decompositions

$$V = V^+ \oplus V^- \oplus V_0 = \widetilde{V}^+ \oplus \widetilde{V}^- \oplus V_0$$

with the properties as indicated above, and we want to prove

$$\dim V^+ = \dim \widetilde{V}^+, \quad \dim V^- = \dim \widetilde{V}^-.$$

We again emphasize that the actual subspaces of V (e.g., V^+ and $\widetilde{V}^+$) can be rather different from each other; we are merely claiming they must have the same dimension.

We will prove $\dim V^+ \leq \dim \widetilde{V}^+$. By symmetry, such a general result with the roles of the decompositions reversed then yields the reverse inequality, whence we'll get equality. Then either applying this argument to $-Q$ (which then interchanges the roles of the plus and minus spaces) or else literally doing the same argument again (with mild changes) yields the corresponding dimension equality for the minus spaces. Thus, we now focus on proving

$$\dim V^+ \leq \dim \widetilde{V}^+.$$

One of the few natural ways to show one vector space has dimension bounded by that of another is to create a linear injection of the first space into the second space. This is what we will do: we shall construct a linear injection $V^+ \hookrightarrow \widetilde{V}^+$ by abstract means.

Consider the composite linear map

$$V^+ \hookrightarrow V = \widetilde{V}^+ \oplus \widetilde{V}^- \oplus V_0 \twoheadrightarrow \widetilde{V}^+.$$

We will prove that this is injective, which will finish the proof (for reasons we have already indicated). Choose $v^+ \in V^+$ in the kernel of this composite map. In other words, relative to the direct sum decomposition

$$V = \widetilde{V}^+ \oplus \widetilde{V}^- \oplus V_0$$

the vector $v^+ \in V$ has vanishing first component, which is to say

$$v^+ = \widetilde{v}^- + v_0$$

for some $\widetilde{v}^- \in \widetilde{V}^-$ and $v_0 \in V_0$. Now since $v^+ \in V^+$, we have $Q(v^+) \geq 0$ with equality if and only if $v^+ = 0$. We will actually prove $Q(v^+) \leq 0$, whence indeed $Q(v^+) = 0$ and so $v^+ = 0$ as desired.

We simply compute

$$\begin{aligned} Q(v^+) = B_Q(v^+, v^+) &= B_Q(\widetilde{v}^- + v_0, \widetilde{v}^- + v_0) \\ &= B_Q(\widetilde{v}^-, \widetilde{v}^-) + B_Q(\widetilde{v}^-, v_0) + B_Q(v_0, \widetilde{v}^-) + B_Q(v_0, v_0) \\ &= B_Q(\widetilde{v}^-, \widetilde{v}^-) + 0 + 0 + 0 \\ &= Q(\widetilde{v}^-) \\ &\leq 0, \end{aligned}$$

as desired. Note that this calculation uses the negative definiteness property of Q on $\widetilde{V}^-$, as well as the fact that $B_Q(v_0, \cdot) = B_Q(\cdot, v_0) = 0$ for all $v_0 \in V_0$. ■