$\boxed{L25}$ Summary of quadratic forms (so far)

as polynomials in n variables with term quadratic

e.g. $q = x^2 - 2xy = q_A(x,y)$ $A = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$

$= (x,y) \cdot A \begin{pmatrix} x \\ y \end{pmatrix}$ corresponding symmetric matrix.

But we consider two quadratic forms the same if they . are related by linear change of variables.

e.g. $q = x^2 - 2xy = (\underbrace{x-y}_{u})^2 - \underbrace{y}_{v}^2$

$= u^2 - v^2$

$u = x - y$
$v = y$

where $x = u+v$, $y = v$ i.e. $\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{P} \begin{pmatrix} u \\ v \end{pmatrix}$

then

$q = q_A(x,y) = q_{A'}(u,v)$

$A' = P^t A P$

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

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

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

we say A' is congruent to A (an equivalence relation)

Theorem 6.7 stated last time says that if $\mathbb{K}$ has characteristic not 2 (ie $1+1 \neq 0$) then any quadratic form q in n variables $x_1 \dots x_n$ can be written as $q = c_1 y_1^2 + \dots + c_n y_n^2$ for a suitable linear substitution to new

<u>Proof</u> We'll proceed by induction on n. Let

$$q = \sum_{i,j=1}^{n} a_{ij} x_i x_j \quad \text{for } A = (a_{ij}) \text{ symmetric}$$

<u>Case 1</u> Assume $a_{ii} \neq 0$ for some i. By permutation (which is a linear substitution) we can assume w.l.o.g. that $a_{11} \neq 0$.

Now let $y_1 = x_1 + \sum_{i=2}^{n} \frac{a_{1i}}{a_{11}} x_i$

Squaring this $\Rightarrow$

$$a_{11} y_1^2 = a_{11} x_1^2 + 2 \sum_{i=2}^{n} a_{1i} x_1 x_i + \underbrace{a_{11} \sum_{i,j=2}^{n} \left(\frac{a_{1i}}{a_{11}}\right) x_i \left(\frac{a_{1j}}{a_{11}}\right) x_j}_{q'(x_2 \dots x_n)}$$

so
$$q = a_{11} x_1^2 + 2 \sum_{i=2}^{n} a_{1i} x_1 x_i + \sum_{i,j=2}^{n} a_{ij} x_i x_j$$

$$= a_{11} y_1^2 + \underbrace{\sum_{i,j=2}^{n} a_{ij} x_i x_j - q'(x_2 \dots x_n)}_{q''(x_2, \dots x_n)}$$

quadratic form in $n-1$ variables

$\therefore$ by induction hypothesis, we assume it can be diagonalized with new variables $y_2 \dots y_n$ in place of $x_1 \dots x_n$

$\therefore q = a_{11} y_1^2 + \sum_{i=2}^{n} c_i y_i^2$ as needed with $c_1 = a_{11}$

Case 2 If all $a_{ii} = 0$, but $a_{ij} \neq 0$ some $i \neq j$

For the pair i, j, write

$$x_i x_j = \frac{1}{4}\left((x_i + x_j)^2 - (x_i - x_j)^2 \right)$$

$$= x_i'^2 - x_j'^2, \qquad x_i' = \frac{x_i + x_j}{2}, \quad x_j' = \frac{x_i - x_j}{2}$$

we make the substitution of x_i, x_j to x_i', x_j' after which q has a non-zero diagonal, so we are now back to case 1.

Case 3 If all $a_{ij} = 0$ then $q = 0$ and statement holds with $c_1 = c_2 \cdots = c_n = 0$ and $y_i = x_i$.

QED.

Example 6.8 Consider the quadratic form

$$q = x^2 + 2xy + 4xz + y^2 + 4z^2$$

as a function (or polynomial) in 3 variables x, y, z over $\mathbb{R}$ (or abstractly $q : V = \mathbb{R}^3 \to \mathbb{R}$ with $v \leftrightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix}$))

To put this in diagonal form we can follow the steps of the proof, or we can do it "by inspection" such as by "completing the square".

$$q = (\underbrace{x + y + 2z}_{u})^2 - 4yz$$

$$= (x + y + 2z)^2 - (\underbrace{y + z}_{w})^2 + (\underbrace{y - z}_{v})^2$$

$$= u^2 + v^2 - w^2$$

is diagonal ~~or~~ in terms of new variables $\begin{bmatrix} u \\ v \\ w \end{bmatrix}$

From the matrix point of view

$$q = q_A(x,y,z) = [x,y,z]\, A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

where $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 4 \end{bmatrix}$ $\quad x^2,\ y^2,\ 2xy,\ 4xz,\ 4z^2$

(dividing by 2 for the off-diagonal entries)

This must be congruent to $A' = P^t A P = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & -1 \end{bmatrix}$

P is the transition matrix to new variables

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

we found
$u = x + y + 2z$
$v = y - z$
$w = y + z$

$y = \frac{v+w}{2}$
$z = \frac{w-v}{2}$

$u = x + \frac{v+w}{2} + w - v$

$\Rightarrow x = u - \frac{3}{2}w + \frac{v}{2}$

$$\Rightarrow P = \begin{bmatrix} 1 & \frac{1}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

We can't say much more in general about the diagonal form (over a general field) but given $q = \alpha_1 x_1^2 + \dots \alpha_n x_n^2$ $\quad \alpha_i \in \mathbb{K}$

<u>if</u> $\sqrt{\alpha_i}$ exist (eg over $\mathbb{K} = \mathbb{C}$) then we

can define $y_i = \begin{cases} \sqrt{\alpha_i}\, x_i & \alpha_i \neq 0 \\ x_i & \alpha_i = 0 \end{cases}$ (say $\alpha_1 \dots \alpha_r \neq 0$)

$\Rightarrow \quad q = y_1^2 + \dots y_r^2$ some $0 \leq r \leq n$

ie $A' = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{array}\right]$ (cf. the canonical form for equivalence)

this is the canonical form for congruence over $\mathbb{C}$.

Similarly over $\mathbb{R}$, suppose $\alpha_1, \dots, \alpha_s > 0$
$\alpha_{s+1} \dots \alpha_{s+t} < 0$
$\alpha_{s+t+1}, \dots, \alpha_n = 0$

and set $y_i = \begin{cases} \sqrt{\alpha_i}\, x_i & 1 \leq i \leq s \\ \sqrt{-\alpha_i}\, x_i & s+1 \leq i \leq s+t \\ x_i & \text{else.} \end{cases}$

$\Rightarrow \quad q = y_1^2 + \dots + y_s^2 - y_{s+1}^2 \dots - y_{s+t}^2$

So $A' = \begin{bmatrix} I_s & & 0 \\ & -I_t & \\ 0 & & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{bmatrix}$ This is the canonical form for congruence over $\mathbb{R}$

we'll see that this does not depend on how we did the diagonalization of an original q. In particular (s, t) are associated to q.

L26

6.3 Abstract quadratic forms and bilinear forms

We already studied linear forms $V \to \mathbb{K}$.
Similarly, we define:

Def. 6.9 (a) A bilinear form on a v.s. V over $\mathbb{K}$ is a function $b: V \times V \to \mathbb{K}$ (ie a function of <u>two</u> vectors) which is linear in each variable with the other variable held fixed:

$$b(v, w_1 + w_2) = b(v, w_1) + b(v, w_2)$$
$$b(v, cw) = c\, b(v, w)$$

$$b(v_1 + v_2, w) = b(v_1, w) + b(v_2, w)$$
$$b(cv, w) = c\, b(v, w)$$

$\forall\ v_1, v_2, w_1, w_2 \in V$, $c \in \mathbb{K}$.

We say b is <u>symmetric</u> if $b(v, w) = b(w, v)$ for all $v, w \in V$.

(b) (with $\mathbb{K}$ not of characteristic 2)

An abstract quadratic form on V is a function $q: V \to \mathbb{K}$ such that

$$q(cv) = c^2 q(v) \qquad \forall\ v \in V,\ c \in \mathbb{K}.$$

<u>and</u> $b(v, w) := \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right)$ $\forall\ v, w \in V$ is a bilinear form on V.

Note that $b(v,w)$ in (6) is automatically symmetric. Moreover, given a symmetric bilinear form $b : V \times V \to \mathbb{K}$ we can set $q(v) = b(v,v)$ and check

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

using the bilinearity of b to expand $b(v+w, v+w)$. So:

$$\boxed{\text{quadratic forms on } V \text{ and symmetric bilinear forms are equivalent data}}$$

(just as quadratic form functions and symmetric matrices were equivalent data).

Example $V = \mathbb{K}$, $v = x$ in a standard basis, $q = x^2$. check:

$$\begin{aligned} b(x,y) &= \tfrac{1}{2}\left(q(x+y) - q(x) - q(y)\right) \\ &= \tfrac{1}{2}\left((x+y)^2 - x^2 - y^2\right) \\ &= xy \qquad \forall\, x,y \in V = \mathbb{K} \end{aligned}$$

is the associated bilinear form.

Example $V = C([0,1])$ continuous functions on $[0,1]$ which are square integrable (one can say better $L^2([0,1])$)

If $f \in V$

$$q(f) = \int_0^1 f(x)^2\, dx$$

is required to exist. The associated bilinear form b is:

$$b(f,g) = \frac{1}{2}\left(\int_0^1 (f+g)^2(x)\,dx - \int_0^1 f^2(x)\,dx - \int_0^1 g^2(x)\,dx\right)$$
$$= \int_0^1 f(x)\,g(x)\,dx$$

$\forall f,g \in V$. This is clearly bilinear in f, g.

Spot quiz Over $\mathbb{R}$, $V = M_2(\mathbb{R})$. Is $q: V \longrightarrow \mathbb{R}$ defined by $q = \det$ a quadratic form? In standard basis of $M_2(\mathbb{R}) \ni v \longleftrightarrow v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $q(v) = ad - bc$.

☐ yes ☑ no

$$b(v,w) = \frac{1}{2}\left(\det(v+w) - \det(v) - \det(w)\right)$$

is not bilinear in v, w.

Example $V = M_2(\mathbb{R})$, $q(v) = \mathrm{Tr}(v^t v)$

So $q: V \to \mathbb{R}$ and $q(cv) = c^2 q(v)$ $\forall c \in \mathbb{R}$

and
$$b(v,w) = \frac{1}{2}\left(\mathrm{Tr}((v+w)^t (v+w)) - \mathrm{Tr}(v^t v) - \mathrm{Tr}(w^t w)\right)$$
$$= \frac{1}{2}\left(\mathrm{Tr}(w^t v) + \mathrm{Tr}(v^t w)\right)$$
$$= \mathrm{Tr}(v^t w) \qquad \left(\text{as } \mathrm{Tr}(w^t v) = \mathrm{Tr}((w^t v)^t) = \mathrm{Tr}(v^t w)\right)$$

is bilinear in v, w.

If we do have a basis $v_1 \ldots v_n \in V$ so that

$$v \longleftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \quad \left(v = \sum_i x_i v_i\right)$$

Then we can view $q(v)$ via this correspon-dence as a function on $\mathbb{K}^n$,

$$q(v) = q_A(x_1 \ldots x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

for $a_{ij} := b(v_i, v_j)$. This is because, using the properties of q.

$$q(v) = q\left(\sum_i x_i v_i\right) = q\Big(x_1 v_1 + \underbrace{\sum_{i=2}^{n} x_i v_i}_{w}\Big)$$

$$= q(x_1 v_1) + q(w) + 2\, b(x_1 v_1, w)$$

$$= x_1^2 \underbrace{q(v_1)}_{b(v_1, v_1)} + 2 \sum_{i=2}^{n} x_1 x_i\, b(v_1, v_i) + q(w)$$

using b bilinear

repeat for w etc (ie prove by induction)

$$\Rightarrow q(v) = x_1^2 a_{11} + 2 \sum_{j=1}^{n} x_1 x_j a_{1j}$$

$$+ x_2^2 a_{22} + 2 \sum_{j=3}^{n} x_2 x_j a_{2j} + q\left(\sum_{j=3}^{n} x_j v_j\right)$$

(etc)

$$= \sum_i x_i^2 a_{ii} + 2 \sum_{i<j} x_i x_j a_{ij}$$

$$= \sum_{i,j=1}^{n} x_i x_j a_{ij}, \quad \text{as claimed.}$$

Summary: fixing a basis gives a quadratic form function in n variables with associated matrix A the values of the associated bilinear b on the basis.

$$A = (a_{ij}) \quad \text{symmetric matrix}$$

$\longrightarrow$ symmetric bilinear form $b(v_i, v_j) = a_{ij}$

$q(v) = b(v,v) \Big\updownarrow$ $b(v,w) = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right)$ "polarization"

$\longrightarrow$ quadratic form functions $q_A(x_1 \ldots x_n) = \sum_{i,j} a_{ij} x_i x_j$

or $q(v)$ abstractly.

(Choosing a different basis leaves $q(v)$, $b(v,w)$ unchanged but the matrix A changes to $P^t A P$ as seen before.

$\boxed{\text{L27}}$ We have a third point of view. Recall V^* is the v.s. of linear maps $V \to \mathbb{K}$.

Let $\alpha : V \longrightarrow V^*$ be a linear map, so

$\alpha(v) \in V^*$, $\alpha(v)(w) \in \mathbb{K}$. Then

$b(v,w) := \alpha(v)(w)$ is a bilinear form.

(it is symmetric iff $\alpha(v)(w) = \alpha(w)(v) \ \forall v, w \in V$)

Conversely, given a bilinear form $b : V \times V \longrightarrow \mathbb{K}$, define $\alpha : V \longrightarrow V^*$ by

$\alpha(v) : V \longrightarrow \mathbb{K}$, defined by $\alpha(v)(w) = b(v,w)$

(symmetric if α obeys the condition above).

Prop TFAE for a v.s. V over $\mathbb{K}$ not of characteristic 2.

(a) quadratic form $q : V \longrightarrow \mathbb{K}$

(b) a symmetric bilinear form $b : V \times V \longrightarrow \mathbb{K}$

(c) a linear map $\alpha : V \longrightarrow V^*$ s.t.
$\alpha(v)(w) = \alpha(w)(v) \quad \forall \; v, w \in V$.

Moreover, if we fix a basis of V then the corresponding matrix A : defined as above for cases (a), (b), and as usual for α as a linear map, are the same. [Here V^* has a dual basis $f_1 \dots f_n \in V^*$ defined by $f_i(v_j) = \delta_{ij}$ if $\{v_i\}$ basis of V].

proof of (c) Given the dual basis as explained on V^*, $\alpha(v_i) = \sum_k a_{ki} f_k$ defines $A = (a_{ij})$ as the corresponding matrix (a special case of our analysis for $\alpha : V \to W$). In our case,

$$\alpha(v_i)(v_j) = \sum_k a_{ki} \underbrace{f_k(v_j)}_{\delta_{kj}} = a_{ji}$$

$$\alpha(v_j)(v_i) = \dots = a_{ij}$$

so the condition on α in part (c) is $\Leftrightarrow a_{ij} = a_{ji}$ as in parts (a), (b). Also $\alpha(v_i)(v_j) = b(v_i, v_j)$ under our correspondence between maps α and bilinear maps, and $= a_{ij}$ as before. QED

Also, if B' is a new basis of V, the matrix A associated to α changes to

$$A' = P_{C,C'}^{-1} \, A \, P_{B,B'}$$

where C is dual basis to the original basis B on V and C' is the dual basis to B'.

Then $P_{C,C'} = (P_{B,B'}^{t})^{-1}$ so

$A' = P^t A P$ as before, where $P = P_{B,B'}$

6.4 Canonical form over $\mathbb{C}$ and $\mathbb{R}$

Theorem 6.12 any $n \times n$ matrix over $\mathbb{C}$ is congruent to a matrix of the form

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \quad r = \text{rank}(A).$$

Moreover, A' is congruent to A iff A' has the same rank as A (This says that the canonical form is unique)

Proof We've seen that over $\mathbb{C}$, $q_A(x_1 \dots x_n)$ can be diagonalized by a linear change of variables given by an invertible matrix P. Here P^t is also invertible (with $(P^t)^{-1} = (P^{-1})^t$). Moreover, rank $(P^t A P)$

$= \operatorname{rank}(A)$ by earlier results. Q.E.D.

Theorem 6.13 (Sylvester's law of inertia). Any $n \times n$ symmetric matrix over $\mathbb{R}$ is congruent to a matrix of the form

$$\begin{bmatrix} I_s & & 0 \\ & -I_t & \\ 0 & & 0 \end{bmatrix}$$

for some associated s, t

Moreover, if A' is congruent to A then it has the same values of s, t (here $s+t$ is the rank and $s-t$ is called the "signature")

Remark the corresponding statement for quadratic forms is that after a linear change of variables, working over $\mathbb{R}$,

$$q(x_1, \ldots x_n) = y_1^2 + \cdots + y_s^2 - y_{s+1}^2 \cdots - y_{s+t}^2$$

with new variables y_i and s, t uniquely determined by the quadratic form (Also called Sylvester's law of inertia).

Proof We've already seen that we can put q into the required form with $\pm 1, 0$ on the diagonal. We also know that $s+t = \operatorname{rank}(A)$ is independent of the linear change of coordinates. Suppose we have two different diagonalizations of A with s, t

and s', t' respectively. (we want to show that $s = s'$, $t = t'$). Wlog suppose $s < s'$ and aim for a contradiction. So

$$q = y_1^2 + \dots + y_s^2 - y_{s+1}^2 \dots - y_{s+t}^2$$
$$= z_1^2 + \dots + z_{s'}^2 - z_{s'+1}^2 - \dots - z^2_{\underbrace{s'+t'}_{s+t}}$$

Now consider the equations

$$y_1 = y_2 = \dots = y_s = 0, \quad z_{s'+1} = \dots = z_n = 0$$

as a system of $s + (n - s') < n$ equations on $x_1 \dots x_n$. This has a <u>nonzero</u> solution (as in the kernel of a linear map on $\mathbb{R}^n$ of rank $< n$). For this value of $x_1, \dots x_n$ we can't have all $y_1 \dots y_n = 0$ or all the $z_1 \dots z_n = 0$ (since a change of variables is invertible)

$$\Rightarrow \quad q = -y_{s+1}^2 \dots - y_{s+t}^2 \quad < 0$$

and

$$= z_1^2 + \dots + z_{s'}^2 \quad > 0$$

which is a contradiction. So $s < s'$ is not possible.

QED.

End with some definitions about quadratic forms q over $\mathbb{R}$.

We say that q (or corresponding b, A) is

- <u>positive definite</u> if $q(v) > 0$ $\forall v \neq 0$, $v \in V$

(ie $b(v,v) > 0$ $\forall v \neq 0$, ie $s = n$, $t = 0$ in Sylvester's law of inertia)

- <u>positive semidefinite</u> if $q(v) \geq 0$ $\forall v \in V$

(ie $b(v,v) \geq 0$ $\forall v$, $t = 0$ in Sylvester's)

- <u>negative definite</u> if $q(v) < 0$ $\forall v \neq 0$

(ie $t = n$, $s = 0$ in Sylvester's)

- <u>negative semidefinite</u> if $q(v) \leq 0$ $\forall v$

($s = 0$ in Sylvester's)

- <u>indefinite</u> otherwise ($s > 0$, $t > 0$ in Sylvester's law of inertia)

$q(v) = b(v,v)$ can be both positive and negative valued as v varies over V.

<u>Reminder</u> UG research seminar today

Learning support now in cafe

Quiz 4 is due Friday by 11.59 pm (a day later since started a day late).