Linear Algebra

July 5, 2008

This course is related to last year's Algebra and Geometry, but takes a more abstract approach. See the website for example sheets.

Books

CW Curtis has written a good book with a title along the lines of "Linear Algebra...", as have K Hoffman and R Kuhze; there are generally a lot of reasonable books on this subject.

Part I
Vector Spaces

We use F to denote the field $\mathbb{R}$ or $\mathbb{C}$; recall that a field F is an abelian group under "+" with identity "0" such that $F \setminus \{0\}$ is an abelian group under "$\times$" which is distributive over "+". The identity for $\times$ is called "1"; $\mathbb{F}_p$ the integers modulo p is a good example of a field

Definition

A vector space V over the field F is a set which forms an abelian group under "+" with identity $\vec{0}$ and is closed under scalar multiplication, which satisfies $\forall v, v_i \in V, \lambda, \lambda_i \in F$ (note nonzero vectors are not underlined in this course):

1. $\lambda (v_1 + v_2) = \lambda v_1 + \lambda v_2$
2. $(\lambda_1 + \lambda_2) v = \lambda_1 v + \lambda_2 v$
3. $(\lambda_1 \lambda_2) v = \lambda_1 (\lambda_2 v)$
4. $1v = v$

This is not the most basic set of axioms, but it is easy enough to check. Technically we should talk about $(V, F, +, \times)$ where the latter two are the vector addition and scalar multiplication operations.

Proposition

If V is a vector space over F then for $\lambda \in F, v \in V$, $\lambda v = \vec{0} \Leftrightarrow \lambda = 0$ or $v = 0$.

Proof

$\vec{0} + 0v = (0+0)\,v = 0v + 0v$ so $0v = \vec{0} \forall v \in V$. $\vec{0} + \lambda\vec{0} = \lambda\left(\vec{0} + \vec{0}\right) = \lambda\vec{0} + \lambda\vec{0}$ so $\lambda\vec{0} = \vec{0} \forall \lambda \in F$.

Now if $\lambda v = \vec{0}$ with $\lambda \neq 0$, $\exists \lambda^{-1} \in F$ so $v = \lambda^{-1}\lambda v = \lambda^{-1}\vec{0} = \vec{0}$.

As an excercise the reader should show that $-1v = v$.

Example

For a set X, the set $F^X = \{f : X \to F\}$ with $(f_1 + f_2)(x)$ defined as $f_1(x) + f_2(x)$ and $(\lambda f)(x)$ defined as $\lambda(f(x))$ is a vector space, as we can prove by checking our definition.

Definition

For V a vec sp over a field F, $U \subset V$, is a subspace, written $U \leq V$, if $\vec{0} \in U$, $u_1 + u_2 \in U$ and $\lambda u \in U$, $\forall u, u_1, u_2 \in U, \lambda \in F$. For example, $\mathbb{R}^{\mathbb{R}}$has a subspace $\mathcal{C}(\mathbb{R})$, the set of cnts real-vald funcs.

Lemma

Any such U forms a vector space over F with the restrictions of $+$ and $\times$ to U

Linear Combinations

The empty lin comb is valid and $\vec{0}$. Take finite; $\sum_{i \in I} \lambda_i v_i$ for an arbitrary indexing set I is only a valid lin comb if all but finitely many of the λ_i are 0.

Spans or generates defined; V is finite dimensional if it is spanned by a finite set. Lin indep defined; v_i for $i \in I$ is lin ind if every finite subcollection therof is. Bases defined; $S \subset V$ is a basis if it spans and is lin ind.

$v_1, \ldots, v_n$ are a basis iff each $v \in V$ has a unique expression in terms of them; if two such expressions for any v when v_i span, difference is $\vec{0}$ so differences of coeffs must be 0 so they are the same; if each v has such an expression, the v_i span and uniqueness means lin ind (else two expressions for $\vec{0}$) so form a basis.

If $v_1, \ldots, v_m$ span, some subset therof is a basis, as if they are lin ind we are done, otherwise we have some l for which $v_l = \alpha_1 v_1 + \cdots + \alpha_{l-1} v_{l-1}$, so can remove it and continue.

The steinitz exchange lemma: given $v_1 \ldots v_n$ lin ind and $w_1 \ldots w_m$ spanning, can replace n of the w_i with v_i and still have them spanning - write v_1 in terms of w_i, rewrite to have one of the w_i in terms of the other w_i and v_1, continue. This implies $n \leq m$.

Main thm: if V is a fin dim vec sp any two bases $v_1, \dots v_n$ and $w_1, \dots, w_m$ have the same number $\dim_F V$ of elts, as the v_i are indep and the w_i span so $n \leq m$ and vice versa, so $n = m$. Note that $\dim_{\mathbb{C}} \mathbb{C} = 1$ but $\dim_{\mathbb{R}} \mathbb{C} = 2$.

An immediate corollary of Steinitz is that if V is a fin dim vec sp over F and $w_1, \dots, w_l$ a lin ind set of vectors of V we can extend it to form a basis. For an n dim vec sp, any lin ind set has $\leq n$ elts with equality only if it is a basis; likewise any spanning set has $\geq n$ elts with equality only if it is a basis. For $v_1, \dots, v_n$ it is equivalent that this is a basis, spanning set or linearly indep. For V a vec sp over F and $S \subset V$ we write $\langle S \rangle$ for the smallest subspace of V which contains S; this is clearly the set of all finite lin combs of elts of S. The intersections of subspaces are subspaces, but their unions almost never; we define for $U, W \leq V$ $U + W = \{u + w : u \in U, w \in W\} = \langle U \cup W \rangle$. This is a subspace, and if U, W fin dim then it is fin dim with dim $\dim U + \dim W - \dim U \cap W$. We prove all these results using bases. $V = U \oplus W$ if every elt of V can be expressed uniquely as $u + w$; W is called the direct complement of U in V. This is equivalent to $V = U + W$ and $U \cap W = \left\{\vec{0}\right\}$ or that for any bases B_1 of U and B_2 of W, $B_1 \cup B_2$ is a basis of V. The second defn implies the first since any $v \in V$ is $u + w$ for some u, w and if $u_1 + w_1 = u_2 + w_2$ then $u_1 - u_2 = w_2 - w_1$; the value for this is $\in U \cap W$ so must be $\vec{0}$ and $u_1 = u_2, w_1 = w_2$. The first implies the third by for B_1 a basis for U, B_2 a basis for W and $B = B_1 \cup B_2$; clearly have B spanning $U + W$, and if $\sum_B \lambda_v v = \vec{0}$ then since representation as $u + v$ is unique, $\sum_{B_1} \lambda_u u = \vec{0}$ and the λ_u are 0, sim the λ_w, so all the λ_v are 0 and B is lin ind. Finally the third implies the second as for $v \in V$ we can express v in B so in B_1 and B_2 so as $u + w$, and if $v \in U \cap W$ we have $v \in U$ so $v = \sum_{B_1} \lambda_u u$ and similarly, so $\sum \lambda_u u - \sum \lambda_w w = \vec{0}$ meaning $\lambda_u \equiv 0$ and similar so $v = \vec{0}$.

Lemma

If V a fin dim vec sp over F and $U \leq V$, $\exists$ a (not generally unique) complement to U - take a basis for U and extend it to one for V and the span of the extension is then such a complement.

Lemma

For $V_1, \dots, V_l \leq V$ with $\sum V_i = \{v_1 + \dots + v_l : v_i \in V_i\}$ this sum is direct if whenever $v_1 + \dots + v_l = v_1' + \dots + v_l'$, $v_i = v_i'$; in this case we write it as $\oplus V_i$; this is equivalent to that $V_i \cap \sum_{j \neq i} V_j = \left\{\vec{0}\right\} \forall i$ or that for any bases B_i of V_i their union $B = \bigcup_i B_i$ is a basis of $\sum V_i$; the reader should proove these equivalences as an exercise.

Quotient Spaces

For V a vec sp over F and $W \leq V$ the quotient group $\frac{V}{W}$ (W is normal since V abelian) is a vec sp over F with addition and scalar multiplication defined in the obvious way. If V is fin dim so is $\frac{V}{W}$; prove this by extending a basis of W.

Part II
Lin Maps and Matricies

1 Defn

For V, W vec sps over F, $\alpha : V \to W$ is linear or a homeomorphism if $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2)$ and $\alpha(\lambda v) = \lambda\alpha(v)$ $(\forall v, v_1, v_2 \in V, \lambda \in F)$.

2 Eg

The function $f \mapsto f'$ on $\mathbb{R}^{\mathbb{R}}$ or $f \mapsto \int_0^x f(t)\,dt$ on $\mathcal{C}[0,1]$, or for any $m \times n$ matrix with entries in F the mapping $\alpha : F^m \to F^n$ $x \mapsto Ax$.

3 L

For U, V, W vec sps over F, the identity and composition of linear maps are linear

4 L

For V, W vec sps over F and α_0 any map of a basis B of V to W, $\exists$ a unique lin map α extending α_0 - proof by basis representation of v and linearity.

5 Note

We often define a lin map just on the basis and then "extend linearly". Also this means if two lin maps between the same spaces agree for a basis of the first space they are equal.

6 Def

A bij lin map is an isomorphism; if $\exists$ one $V \to W$ we write $V \simeq W$.

7 L

$\simeq$ is an equiv rel on the set of all vec sps over F; only hard part of proof is linearity of α^{-1}, which must exist as α bij. For $w_1, w_2 \in W$, write the w_i as

$\alpha^{-1}v_i$, then $\alpha^{-1}(w_1+w_2)=\alpha^{-1}(\alpha v_1+\alpha v_2)=\alpha^{-1}\alpha(v_1+v_2)$ by linearity of α; rest of proof similar.

8 Thm

If V a vec sp of dim n over F then $V\simeq F^n$. Express vecs of v in terms of basis and then map to F^n in the obvious way

9 Thm

The vec sps U, W over F are isom if they have the same dim - obvious corollary. The converse is also true; for $\alpha: U\to W$ an isomorphism and B a basis for U, $\alpha(U)$ is a basis for W - spans since B spans U and α surj, sim lin ind.

10 Def

For $\alpha: V\to W$ linear, the nullity $N(\alpha)=\{v\in V:\alpha(v)=0\}$, also sometimes $\ker\alpha$; sim $Im(\alpha)=\{w\in W: w=\alpha(v) \text{ for some } v\in V\}$; note the former is a subspace of V and the latter of W. α is inj iff $\ker\alpha=\left\{\vec{0}\right\}$, surj iff $Im\alpha=W$; we define the rank $rk(\alpha)$ or $r(\alpha)$ by $\dim Im\alpha$, nullity $n\alpha=\dim N\alpha$. [missing brackets because I'm cool]

11 Rank-Nullity Thm

For V, W vec sps over F with $\dim_F V$ fin, $\alpha: V\to W$ linear $\dim V=r\alpha+n\alpha$; take a basis for $N\alpha$, extend this to a basis of V and the image of the extension is a basis for $Im\alpha$, or can prove by iso from $\frac{V}{N\alpha}$ to $Im\alpha$.

12 L

for V, W vec sps over F of equal fin dim and $\alpha: V\to W$ linear, equivalent that:

1. α iso
2. α inj
3. α surj

Proove by rank-nullity

13 Prop: the space $L(V, W)$ of linear maps $V \to W$ for V, W vec sps over F is a vec sp

Also sometimes called $\mathrm{Hom}(V, W)$; vec sp under addition and multiplication defined pointwise. If V, W fin dim so is L, with dimension $\dim V \dim W$; proof of this later (19)

Matricies

An $m \times n$ matrix A over F is an array with m rows and n columns with entries $\in F$, we usually write them as (a_{ij}) with individual elts a_{ij}; the set of all such is $M_{m,n}(F)$.

14 Prop

This is a vec sp over F with addition and multiplication defined pointwise; dimension $m \times n$ by the standard basis (the set of matricies with 0s in all but one entry, which contains a 1)

Repr of lin maps by matricies

For V, W fin dim vec sps over F and $\alpha : V \to W$ linear fix bases $B = \{v_1, \ldots, v_n\}, C = \{w_1, \ldots, w_m\}$, then for $v = \sum \lambda_i v_i \in V$ write $[v]_B = \begin{pmatrix} \lambda_1 \\ \ldots \\ \lambda_n \end{pmatrix}$, sim $[W]_C$.

15 Defn

$[\alpha]_{B,C}$ the matrix of α wrt B, C is $([\alpha v_1]_C \ldots [\alpha v_n]_C)$. [The lecturer has the dimensions of his matricies hopelessly confused, so I'm ignoring them].

16 L

$\forall v \in V$, $[\alpha v]_C = [\alpha]_{B,C} [v]_B$ multiplied as matricies.

17 Rk

We get the same result by mapping $v \in V$ to a vector $w \in W$ by α and then representing this as a column in F^m as by mapping v to a column in F^n and then applying the corresponding matrix A.

18 Rk

This matrix $[\alpha]_{BC}$ is the only matrix A for which $[\alpha v]_C = A[v]_B \forall v \in V$ by taking v to be the basis vectors of V.

19 Prop

For V, W vec sps over F with dim n, m respectively $L(V, W) \simeq M_{m,n}(F)$; fix bases and then map $\alpha \mapsto [\alpha]_{BC}$; inj as if mapping is 0 α is 0 on a basis so the 0 map, surj as let α map the bases as indicated by a given matrix and extend linearly; this prooves 13 above.

20 L

For $\beta : U \to V$ and $\alpha : V \to W$ linear, for bases A, B, C respectively of U, V, W, $[\alpha \circ \beta]_{AC} = [\alpha]_{BC}[\beta]_{AB}$ by action on basis vectors of U.

Change of Bases

For bases $B = v_1, \ldots, v_n, B'$ of a vec sp V the matrix $P = p_{ij}$ given by $v'_j = \sum_i p_{ij} v_i$ is the change of basis matrix from B to B'; it looks like $([v'_1]_B \ldots [v'_n]_B)$ and we can see it as $[i]_{B'B}$. Then we have $[v]_B = P[v]_{B'}$ either by 16 or directly by actuan on basis vectors. Note P must be invetible since P^{-1} is the change of basis matrix from B' to B, $[i]_{BB'}$; $[i]_{B'B}[i]_{BB'} = [i]_{BB} = I$ and similarly the product in the other direction.

L

For $\alpha : V \to W$ linear $A = [\alpha]_{BC}$ and $A' = [\alpha]_{B'C'}$, $A' = Q^{-1}AP$ for some invertible Q, P as $Q[\alpha V]_{C'} = [\alpha v]_C = A[v]_B = AP[v]_{B'}$ [I'm guessing what the lecturer meant here] for Q and P the change of basis matricies between B, B' in V and C, C' in W respectively.

Def

The matricies A, A' are equiv if $A' = Q^{-1}AP$ for some invertible Q, P; this clearly defines an equiv rel on $M_{m,n}(F)$

L

1. For V, W vec sps over F of respective dim n, m and $\alpha : V \to W$ linear $\exists$ bases B of V, C of W (not generally unique) st $[\alpha]_{BC} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

for these entries submatricies where I_r is the $r \times r$ identity; compare with rank-nullity, take a basis of V containing a basis for $N\alpha$, then extend its image to a basis for W and done (modulo re-ordering basis vectors)

2. Any matrix is equiv to one of this form

Def

For $A \in M_{m,n}(F)$ the (col) rank of A $r(A)$ is the dim of the subspace of F^m generated by the cols of A; if $A = [\alpha]_{BC}$ this is $r\alpha$ as we have an iso from $Im\alpha$ to the span of the cols of A by $\alpha v \mapsto [\alpha v]_C$

T

The matricies A, A' equiv iff $rA = rA'$; forward implication since both can represent the same lin trans, reverse by A equiv to some $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where $r = rA$ by first part, sim for A' and these are only equiv if $rA = rA'$. Row rank (dim of the span of the rows) of any A is the same as col rank; take A equiv to $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ and then A^T equiv to $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where the 0s may be differently sized and these clearly have the same col rank (rowrk$A$ =colrkA^T) so done.

Eltary ops, Eltary matricies

Def eltary col ops on a matrix A are swap two cols i, j , replace col i by $\lambda\times$ itself, or add $\lambda\times$ col i to col j and sim eltary row ops; these are all reversible. We find the corresponding eltary matricies by performing these operations on I, e.g. in 2×2 $T_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, M_{1\lambda} = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, C_{12\lambda} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$ are matricies of the 3 types; an eltary operation can be performed by postmultiplying A by the corresponding eltary matrix (or premultiplying for a row op), e.g. $A \mapsto AT_{ij}$. We can use this to constructively prove that any matrix is equiv to one of the form $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$; if A has no nonzero entries we stop, otherwise take some $a_{ij} = \lambda \neq 0$, swap rows i1 and cols ij and multiply col 1 by λ^{-1}, then clear out the first row and col by ops of typ. 3 and recurse on the $m-1 \times n-1$ submatrix in the bottom right corner; since every operation can be represented by a matrix we can find P, Q by multiplying the matricies corresponding to the ops we have performed in the correct order. [I assume; I was out trying to kill people for the end of this lecture].

Variations

We need only eltary row ops to obtain the row echelon form of a matrix (use Gaussian elimination)

For A square $(n = m)$ if A non-singular can obtain I_n w/ just eltary col ops (or row ops); inductively if we have the first k rows we have some $j > k$ w/ $a_{k+1j} = \lambda \neq 0$ since otherwise A would be singular [?], so we swap cols $k+1, j$, divide col $k+1$ by λ, and then clear out the remainder of row $k+1$ by type 3 ops. We can use this to constrct A^{-1} by $AE_1 \dots E_C = I_n$ so $I_n E_1 \dots E_C = A^{-1}$. [?]

P 34

Any invertible $n \times n$ mat is a prod of eltary mats - construct from A^{-1} as above.

For $V = W, C = B$ we write $L(V)$ rather than $L(V, V)$, $[\alpha]_B$ rather than $[\alpha]_{BB}$, and $M_n(F)$ for $M_{n,n}(F)$.

D

A, A' are similar or conjugate if $A' = P^{-1}AP$ some invertible P; note $[\alpha]_{B'} = P^{-1}[\alpha]_B P$ for P change of basis mat from B to B'

Det and Trace

Trace

Defined; note linear $M_n F \to F$.

L

$trAB = trBA$

L

Similar mats have same tr as $trP^{-1}AP = trAPP^{-1} = trA$.

For α linear def $tr\alpha = tr[\alpha]_B$ for any basis B; now know well defd.

Recall S_n is the group of permutations of $\{1, \dots, n\}$; let $\epsilon(\sigma) = +$ for σ even, $-$ for σ odd (i.e. the composition of an even or odd no. of transpositions; recall this is well defined. Def $\det A$ by $\sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$. This is the sum of $n!$ summands, each of which is a sign $\times$ a prod of n factors, one from each row and one from each col. Note this is the familiar determinant for e.g. $n = 2$.

We write $A = \left(A^{(1)} A^{(2)} \dots A^{(n)}\right)$ an n-tuple of vectors in F^n; note $I = (e_1, \dots, e_n)$.

Def

The func $d : F^n \times \cdots \times F^n \to F$ is a volume func on F^n if it is multilinear (linear in each argument) and alternating (0 if any two distinct args are the same). d is a determinant form if also $d(e_1, \dots, e_n) = 1$.

L

For d a vol func swapping cols changes sign, as d is linear in both of these cols so $d(a, b, \dots) + d(b, a, \dots) = d(a + b, a + b, \dots) = 0$ and similar.

Corollary

If d a vol func on F^n and $\sigma \in S_n$ $d(v_{\sigma 1} \dots v_{\sigma n}) = \epsilon(\sigma) d(v_1, \dots, v_n)$. In particular for a det form $d(e_{\sigma 1}, \dots, e_{\sigma n}) = \epsilon(\sigma)$.

T

If d a vol func on F^n and $A = \left(A^{(1)} \dots A^{(n)}\right)$, $d\left(A^{(1)}, \dots, A^{(n)}\right) = \det A d(e_1 \dots, e_n)$ which of course $= \det A$ if d is a det form, as it $= d\left(\sum_{j_1=1}^n a_{j_1 1} e_{j_1}, \dots\right) = \sum_{j_1=1}^n a_{j_1 1} d\left(e_{j_1}, A^{(2)}, \dots\right) = \cdots = \left(\prod_{i=1}^n \sum_{j_i=1}^n a_{j_i i}\right) d\left(e_{j_1}, \dots, e_{j_n}\right)$ (by which I mean all the sums are applied); the terms where the j_i are not all distinct are 0 so this is $\sum_{\sigma \in S_n} a_{\sigma(1)1} \dots a_{\sigma(n)n} \epsilon(\sigma)$ as required. This means a det function is unique if it exists

T10

$d : F^n \times \cdots \times F^n \to F$ $\left(A^{(1)}, \dots, A^{(n)}\right) \mapsto \det A$ is a det func on F^n; multilinear as $\det A$ is a sum with each of the summands $\epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$ linear in each factor, alternating as if $A^{(k)} = A^{(l)}$ for some $k \neq l$, let $\tau = (kl) \in S_n$, then we can express the sum $\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$ as $\sum_{\sigma \in A_n} \epsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} + \epsilon(\sigma\tau) a_{\sigma\tau(1)1} \dots a_{\sigma\tau(n)n}$ (where A_n is the alternating group $\sigma \in S_n : \epsilon(\sigma) = +$) which is $\sum_{\sigma \in A_n} a_{\sigma(1)1} \dots a_{\sigma(n)n} - a_{\sigma\tau(1)1} \dots a_{\sigma\tau(n)n}$ but for any $\sigma \in S_n$, $a_{\sigma(1)1} \dots a_{\sigma(n)n} = a_{\sigma\tau(1)1} \dots a_{\sigma\tau(n)n}$ as $k = l$ so this is 0. Finally $\det I = \sum_{\sigma \in S_n} \epsilon(\sigma) e_{\sigma(1)1} \dots e_{\sigma(n)n} = \sum_{\sigma \in S_n} \epsilon(\sigma) \delta_{\sigma(1)1} \dots \delta_{\sigma(n)n}$; the only nonzero summand is where $\sigma = \iota$; $\epsilon(\iota) = +$ so $\det I = 1$. Therefore det is the unique determinant form.

L11

$\det A^T = \det A$ as if $\sigma \in S_n$ then $a_{\sigma(1)1} \dots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \dots a_{n\sigma^{-1}(n)}$, since the same factors are present in both products. We have $\epsilon\left(\sigma^{-1}\right) = \epsilon(\sigma)$ and σ^{-1} runs over S_n as σ does, so replacing σ^{-1} with π, $\det A = \sum_{\pi \in S_n} \epsilon(\pi) a_{1\pi(1)} \dots a_{n\pi(n)} = \det A^T$.

L12

det is the unique multilinear alternating function of rows normalized at I - immediate corollary.

L13

If $A = (a_{ij})$ an upper triangular matrix (i.e. $a_{ij} = 0 \ \forall i > j$) $\det A = a_{11} \dots a_{nn}$ (and similarly the same result for a lower triangular matrix) as for $a_{\sigma(1)1} \dots a_{\sigma(n)n}$ to be nonzero we must have $\sigma(1) \leq 1$ so $\sigma(1) = 1$, then need $\sigma(2) \leq 2$ so $\sigma(2) = 2$ and so on, so the only σ with this nonzero is ι and $\det A = \epsilon(\iota) a_{11} \dots a_{nn} = a_{11} \dots a_{nn}$.

L14

If E an eltary $n \times n$ mat for any A $\det AE = \det A \det E = \det EA$ so performing an eltary op on A multiplies $\det A$ by the det of the corresponding eltary mat,; $\det T_{ij} = -1$ by alternating and applying the transposition multiplies $\det A$ by -1 by the same; $\det M_{i\lambda} = \lambda$ by multilinearity and applying the multiplication multiplies $\det A$ by λ by the same, and $\det C_{ij\lambda} = 1$ since this is upper or lower triangular and the reader should proove the corresponding operation leaves $\det A$ unchanged [since the lecturer apparently can't].

T15

Let A be a square matrix, then A is non-singular iff $\det A \neq 0$; if it is non-singular A is a prod of eltary matricies so has det the product of their dets $\neq 0$ by above; if A is singular we can obtain a matrix w/ a 0 col (since a 0 col is a non-trivial lin comb of the cols of A) by eltary col ops, so det of this matrix is 0 and $\det A = 0$ by above.

T16

For $A, B \in M_n(F)$ $\det AB = \det A \det B$; if B singular so is AB by considering the corresponding lin maps, so $\det AB = 0 = \det A \det B$, otherwise express B as a prod of eltary matricies and $\det AB = \det AE_1 \dots E_C = \det AE_1 \dots E_{C-1} \det E_C = \dots = \det A \det E_1 \dots \det E_C = \det A \det B$.

C17

A invertible $\Rightarrow \det A = \frac{1}{\det A^{-1}}$.

C18

Conjugate mats have same ded as $\det PAP^{-1} = \det A \det P \det P^{-1} = \det A$.

D19

$\det \alpha = \det [\alpha]_B$ for any basis B; well defd by above.

T20

$\det : L(V) \to F$ has $\det \iota = 1, \det \alpha\beta = \det \alpha \det \beta$ and $\det \alpha \neq 0$ iff α non-singular, and $\det \alpha^{-1} = (\det \alpha)^{-1} \,\forall$ such α - from matrix properties.

Rks

$GL(V)$ is the group of all automorphisms of V; an $\alpha \in L(V)$ is an endomorphism and a non-singular (bijective) endomorphism is an automorphism. Say V is n-dim over F; then $GL_n(F)$ is the group of invertible $n \times n$ mats on F and $\det : GL_n(F) \to F$ is a group hom and surj; $\ker \det$ is called $SL_n(F)$, the group of mats w/ det 1. For A $n \times n$ mat representing $\alpha \in L(V)$, equivalent that A non-singular, α non-singular, A invertible, α invertible, $\det A \neq 0$ or $\det \alpha \neq 0$.

L21

For $A \in M_m(F), B \in M_k(F), C \in M_{m,k}(F)$, $\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \det B$ as if we write $n = m+k, X = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ we have $\det X = \sum_{\sigma \in S_n} \epsilon(\sigma)\, x_{\sigma(1)1} \ldots x_{\sigma(n)n}$ but $x_{\sigma(j)j} = 0$ for $j \leq m$ and $\sigma(j) > m$, so we sum only over σ which map $[1,m]$ and $[m+1,n]$ to themselzen, which is the same as summing over $\sigma_1 \in S_m, \sigma_2 \in S_k$ where $\sigma_1(l) = \sigma(l), \sigma_2(l) = \sigma(m+l) - m$; we have $\epsilon(\sigma) = \epsilon(\sigma_1)\epsilon(\sigma_2)$ so $\det X = \left(\sum_{\sigma_1 \in S_m} \epsilon(\sigma_1)\, a_{\sigma_1(1)1} \ldots a_{\sigma_1(m)m}\right)\left(\sum_{\sigma_2 \in S_k} \epsilon(\sigma_2)\, b_{\sigma_2(1)1} \ldots b_{\sigma_2(k)k}\right)$ and done.

L22

Let $A \in M_n(F)$, $A = (a_{ij})$; write $A_{\widehat{ij}}$ for the $n-1 \times n-1$ mat obtained by deleting row i, col j from A. For fixed j, $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ji} \det A_{\widehat{ij}}$; this is the expansion in col j (and by transpose, for fixed i $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{\widehat{ij}}$); note we can define det inductively by this (set $\det(a) = a$ for 1×1 matricies as the base case); $\det A = \det\left(A^{(1)}, \ldots, \sum_{i=1}^{n} a_{ij} e_i, \ldots, A^{(n)}\right) = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det \begin{pmatrix} 1 & \ldots \\ 0 & A_{\widehat{ij}} \end{pmatrix}$ (by repeated col transpositions) $= \sum_{i=1}^{j} (-1)^{i+j} \det A_{\widehat{ij}}$.

D23 [?]

The adjugate or classical adjoint matrix $adj A$ of $A \in M_n(F)$ is the matrix w/ ij entry $(-1)^{i+j} \det A_{\widehat{ji}}$.

T24

$AdjA \times A = (\det A) I$ (as a corollary, if A invertible $A^{-1} = \frac{1}{\det A} adjA$ [no proof it is inverse from both sides here, but maybe we know already?]) since $\det A = \sum_i (adjA)_{ji} a_{ij}$ which is the jj entry of $adjA \times A$; $0 = \det \left(A^{(1)}, \dots, A^{(k)}, \dots, A^{(k)}, \dots, A_n\right) = \sum_i (adjA)_{ji} a_{ik}$ which is the jk entry of $adjA \times A$.

Digression: systems of linear eqns

$A\vec{x} = \vec{b}$ for A a $m \times n$ mat, $\vec{b}$ m-vec and $\vec{x}$ n-vec is a system of m eqns for n unknowns. Recall it has a sol iff $r(A) = r(A \mid b)$ (the augmented matrix formed by A with the extra col $\vec{b}$); if we have such a sol $\vec{b}$ is lin dep on the cols of A and vice versa.

The sol is unique iff this rank $= n$; to find it we use eltary row ops to perform Gaussian elimination.

For $m = n$ and A non-singular the unique sol is $\vec{x} = A^{-1}\vec{b}$.

L25 (Cramer Rule) [sp?]

If A is a non-singular $n \times n$ mat the system $A\vec{x} = \vec{b}$ has $x_i = \frac{\det A_{\hat{i}} b}{\det A} \forall i$ as its unique sol where $A_{\hat{i}}b$ is A with col i replaced by $\vec{b}$, as if $A\vec{x} = \vec{b}$ then $\det A_{\hat{i}}b = \det\left(A^{(1)}, \dots, A^{(i-1)}, \vec{b}, A^{(i+1)}, \dots, a^{(n)}\right) = \det\left(A^{(1)}, \dots, A^{(i-1)}, \sum_j A^{(j)} x_j, A^{(i+1)}, \dots, a^{(n)}\right) = \sum_j x_j \det\left(A^{(1)}, \dots, A^{(i-1)}, A^{(j)}, A^{(i+1)}, \dots, a^{(n)}\right) = x_i \det A$.

C26

If $A \in M_n(\mathbb{Z})$ with $\det A = \pm 1$ and $\vec{b} \in \mathbb{Z}^n$ we can solve $A\vec{x} = \vec{b}$ over $\mathbb{Z}$ [why?]

4 Endomorphisms, mats and evecs

For this section: V is a fin dim vec sp over F, $\dim V = n, B = \{v_1, \dots, v_n\}$ is a basis and $\alpha : V \to V$ is linear so an endomorphism.

We want to pick B st $[\alpha]_B$ is simple; recall $[\alpha]_{B'} = P^{-1} [\alpha]_B P$ for P a change of basis mat from B to B', so equiv that for $A \in M_n(F)$ we want A' conj to A with a nice form.

Def 1

α is diagonalizable if $\exists B : [\alpha]_B$ diagonal, trianglizable if $\exists B : [\alpha]_B$ upper triangular (we could equally well use lower triangular, but must define one or the other, not both. The defns for a matrix A are obvious.

D2

$\lambda \in F$ is an eval if $\exists v \neq 0 \in V : \alpha(v) = \lambda v$.

Rk3

λ is an eval of α iff $\alpha - \lambda\iota$ singular iff λ a root of $\chi_\alpha(t) = \det(\alpha - t\iota)$.

Def $v_\lambda = \{v \in V : \alpha(v) = \lambda v\}$ the $\lambda-$eigenspace of α.

Note 4

The col j of $[\alpha]_B$ is λe_j iff $\alpha(v_j) = \lambda v_j$; $[\alpha]_B$ is diaconal iff B consists of evecs, upper triangular iff $\alpha(v_j) \in \langle v_1, \ldots, v_j \rangle$; note this means v_1 is an evec.

R5

Recall: a func $f : F \to F$ is a polynomial func if it is of the form $f(t) = a_n t^n + \cdots + a_0$ for $n \in \mathbb{N}_0, a_i \in F \forall i$; the largest $m : a_m \neq 0$ is the degree of f with the degree of 0 taken to be $-\infty$; this gives us that $\deg fg = \deg f + \deg g$ (addition and multiplication of polys is defd the obvious way); the polys form a ring $F[t]$. λ is a root of the poly f if $f(\lambda) = 0$; if λ is a root of f then $(t - \lambda)$ divides $f(t)$, as then $f(t) = f(t) - f(\lambda) = a_n(t^n - \lambda^n) + \cdots + a_1(t - \lambda) = (t - \lambda)\left(a_n\left(t^{n-1} + t^{n-2}\lambda + \cdots + \lambda^{n-1}\right) + \cdots + a_1\right) = (t - \lambda) q(t)$ for some $q(t) \in F[t]$; we say λ is a root of f w/ multiplicity e if $(t - \lambda)^e$ divides f but $(t - \lambda)^{e+1}$ does not.

L7 [ya rly]

A poly over F of deg $n \geq 0$ has at most n roots (counted w/ multiplicity); trivially true for $n > 0$, then strong induction; for f a poly of deg $n > 0$ if no roots then done, otherwise let λ a root of multiplicity $e \geq 1$, then $f(t) = (t - \lambda)^e q(t)$ for q a poly of deg $n - e$ over f and any root of $f \neq \lambda$ is a root of q.

C8

If f_1, f_2 polys of deg $< n$ and $f_1(t_i) = f_2(t_i)$ for n points t_i of F then $f_1 = f_2$ by considering $f_1 - f_2$.

Rk9 FTA

Any poly over $F = \mathbb{C}$ of deg $n > 0$ has a root (and so inductively has n roots, counted as always with multiplicity); $\mathbb{C}$ is algebraically closed.

Def

The char poly $\chi_\alpha(t) = \det(\alpha - t\iota)$ (and sim χ_A) is a poly of deg $n \in F[t]$

Rk11

Conj mats have the same char poly (consider corresponding α)

T12

For $F = \mathbb{C}[\alpha]_B$ is upper triangular for some B (so any squar cplx mat is triangable): we induct on n, the $n = 1$ case being trivial. If true $\forall V$ of dim $< n$ for some $n > 1$ any α has some eval λ by FTA, so $\alpha - \lambda\iota$ is singular; put $U = Im\,(\alpha - \lambda\iota) \subsetneqq V$, then U is α-invariant ($\alpha\,(U) \subset U$); consider $\alpha_1 = \alpha\mid_U$; by the induction hypothesis $\exists$ a basis B_1 of U w/ $[\alpha_1]_{B_1}$ upper triangular; extend to B with $\{v_1, \ldots, v_k\} = B_1$. Then $[\alpha]_B = \begin{pmatrix} [\alpha_1]_{B_1} & \star \\ 0 & \lambda I \end{pmatrix}$ (where $\star$ is some matrix) since for $1 \le j \le k$ $\alpha\,(v_j) = \alpha_1\,(v_j)$, so the left hand portion is as given, and for $k < j \le n$ $\alpha\,(v_j) = u_j + \lambda v_j$ for some $u_j \in U$ since $(\alpha - \lambda\iota)\,(v_j) \in U$, so the right hand portion is as given, and the matrix is upper triangular.

Rk13

This is not true for $F = \mathbb{R}$ by e.g. $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

T14

α is triangable iff χ_α can be written as a prod of lin factors, i.e. all evals are $\in F$; necessity by if $[\alpha]_B = A$ upper triangular $\det\alpha = a_{11}\ldots a_{nn}$ and $\chi_\alpha\,(t) = (a_{11} - t)\ldots(a_{nn} - t)$, sufficiency by proof as above; for the inductive step we have $\chi_\alpha\,(t) = \chi_{\alpha_1}\,(t)\,(\lambda - t)^m$ where $m = n - k$.

We can also proove 12 using eigenspaces.

T15

α is diagable if $p\,(\alpha) = 0$ (the zero endomorphism) for some poly p the prod of distinct lin factors; forward implication by let $\lambda_1, \ldots, \lambda_k$ the distinct evals of α which are the nonzero values in $[\alpha]_B = A$ diagonal, then take $p\,(t) = (t - \lambda_1)\ldots(t - \lambda_n)$ and $p\,(\alpha) = 0$ (if $v \in B$, $\alpha\,(v) = \lambda_l v$ some $1 \le l \le k$ so $(\alpha - \lambda_l\iota)\,(v) = 0$ so $p\,(\alpha)\,(v) = 0$ so $p\,(\alpha)$maps B to 0 so is the 0 endomorphism), reverse by if $p\,(t) = (t - \lambda_1)\ldots(t - \lambda_k)$ w/ all the λ_i distinct set $p_j\,(t) = (t - \lambda_1)\ldots(t - \lambda_{j-1})\,(t - \lambda_{j+1})\ldots(t - \lambda_k)$ and $h_j\,(t) = (p_j\,(\lambda_j))^{-1}\,p_j\,(t)$, then $h_j\,(\lambda_l) = \delta_{jl}$; write $h\,(t) = \sum_{j=1}^k h_j\,(t)$ and h is the poly 1 since $h\,(t) - 1$ is a poly of deg $< k$ (since a sum of polys of deg $k - 1$) w/ k roots $\lambda_1, \ldots \lambda_n$; put $\pi_j = h_j\,(\alpha)$, then $\iota = \pi_1 + \cdots + \pi_k$ and $\pi_j\pi_l = 0$ if $j \ne l$ since $p \mid h_j h_l$ and $p\,(\alpha) = 0$; $\pi_j^2 = \pi_j$ since $= \pi_j \sum_l \pi_l$. Put $V_j = Im\,(\pi_j)$, then $V_j \subset v_{\lambda_j}$ since $(\alpha - \lambda_j\iota)\,\pi_j = p\,(\alpha) = 0$; note π_j restricted to V_l is 0 for $j \ne l$, ι_{V_j} for $j = l$ (since $\pi_j\,(\pi_j\,(v)) = \pi_j\,(v)$); now $V = \bigoplus_j V_j$; $V = \sum_j V_j$ since for $v \in V$ $v = \iota\,(v) = \sum_j \pi_j\,(v)$ and if $u_1 + \cdots + u_k = u_1' + \cdots + u_k'$ with $u_j, u_j' \in v_j$ then applying π_j have $u_j = u_j'$ for each j; if B_j is a basis of V_j then the union $B = \bigcup_j B_j$ is a basis of V (spans clearly, lin ind as if $\sum_{v\in B} \lambda_v v = 0$ $\sum_j \left(\sum_{v \in B_j} \lambda_v v\right) = 0$ so $\sum_{v\in B_j} \lambda_v v = 0 \forall j$ so $\lambda_v = 0 \forall v \in B_j \forall j$ and done.

Rk

1. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagable as its evals are 1 but the only mat conj to I is I ($PIP^{-1} = I \forall P$).

2. For $\lambda_1, \ldots, \lambda_k$ the distinct evals of α $\sum_j V_{\lambda_j}$ is direct so the only way diagonalization fails is if $\sum_j V_{\lambda_j} \subsetneq V$ since if $v \in V_{\lambda_j} \cap \sum_{i \neq j} V_{\lambda_i}$ we apply $(\alpha - \lambda_1 \iota) \ldots (\alpha - \lambda_{j-1} \iota)(\alpha - \lambda_{j+1} \iota) \ldots (\alpha - \lambda_k \iota)$ which maps all vecs $\in \sum_{i \neq j} V_{\lambda_i}$ to 0 but multiplies any vec $\in V_{\lambda_j}$ by the non-zero scalar $(\lambda_j - \lambda_1) \ldots (\lambda_j - \lambda_{j-1})(\lambda_j - \lambda_{j+1}) \ldots (\lambda_j - \lambda_k)$, so $v = 0$.

T17

Simultaneous diagation: let α_1, α_2 commuting (necessary as diag mats commute) diagable endomorphisms of V, then they are simultaneously diagable, i.e. $\exists B : [\alpha_1]_B, [\alpha_2]_B$ diagonal; we have $V = V_1 \oplus \cdots \oplus V_k$ where the V_i are the eigensps of α_1; say $\alpha_1(v) = \lambda_j v$ for $v \in V_j$. Then $\alpha_2(V_j) \subset V_j$ as if $v \in v_j$ $\alpha_1(\alpha_2(v)) = \alpha_2(\lambda_j v) = \lambda_j \alpha_2(v)$; now $\alpha_2|_{V_j}$ is diagable by T15 so $\exists$ a basis B_j consisting of evecs of α_2 (which will be evecs of α_1 as well) and $B = \bigcup_j B_j$ is a basis of V consisting of evecs of both α_1 and α_2.

18 Polys over F

Given polys a, b over F w/ $b \neq 0 \exists$ polys q, r w/ $a = bq + r, \deg r < \deg b$ (hence $F[t]$ is a euclidean domain; this has nice consequences, see the IA course N&S); proof inducting by dividing in the obivous way

D19

The min poly m_α of α is the monic (leading coeff 1) poly of smallest deg w/ $m_\alpha(\alpha) = 0$; exists since have a poly of deg $\leq n^2$ w/ $p(\alpha) = 0$ as $\dim_F L(V) = n^2$ so $\iota, \alpha, \ldots, \alpha^{n^2}$ lin dep, unique by:

L21

if $p(\alpha) = 0, m_\alpha \mid p$; write $p = q m_\alpha + r$, then $\deg r < \deg m_\alpha$ but $r(\alpha) = 0$.

T22 - Cayley-Hamilton

$\chi_\alpha(\alpha) = 0$ (and sim for mats); a corollary of this (C23) is that $m_\alpha \mid \chi_\alpha$. For $A \in M_n(F)$ let $(-1)^n \chi_A(t) = t^n + a_{n-1} t^{n-1} + \cdots + a_0 = \det(tI - A)$, now for any mat B $B \times adj B = I \det B$ so we this to $tI - A$; $adj(tI - A)$ is a mat w/ entries polys of deg $\leq n-1$ so we can consider this as a poly w/ mat coeffs $(tI - A)\, adj(tI - A) = (tI - A)\left(B_{n-1} t^{n-1} + \cdots + B_0\right)$ for some mats B_i; comparing coeffs we have $I = B_{n-1}, a_{n-1} I = B_{n-2} - A B_{n-1}, \ldots, a_0 I = -A B_0$;

multiplying the ith eqn by A^{n-i+1} from the left and summing all the eqns we have $A^n + a_{n-1}A^{n-1} + \cdots + a_0 I = 0$. The schedules require only a proof of this result over $\mathbb{C}$, which can be done by other means [it was given in lectures, but I prefer this one].

D24

λ an eval of α has $\chi_\alpha(t) = (t-\lambda)^{a_\lambda} q(t)$ with q a poly not divisible by $t-\lambda$; call the a_λ such that this is the case the algebraic multiplicity of λ (as an eval of α). We def g_λ the geometric multiplicity of λ by $\dim N(\alpha - \lambda\iota)$.

L25

For λ an eval $1 \le g_\lambda \le a_\lambda$; $1 \le g_\lambda$ since $\alpha - \lambda\iota$ singular, $g_\lambda \le a_\lambda$ since for $B = v_1, \ldots, v_g, \ldots, v_n$ containing a basis $v_1, \ldots, v_g$ of $N(\alpha - \lambda\iota)$, $\alpha_B = \begin{pmatrix} \lambda I_g & \star \\ 0 & A_1 \end{pmatrix}$ (I_g being the $g \times g$ identity) for some mat A_1 so $\chi_\alpha(t) = (\lambda - t)^g \chi_{A_1}(t)$.

Now taking $F = \mathbb{C}$:

26

$\chi_\alpha(t) = (\lambda_1 - t)^{a_1} \ldots (\lambda_k - t)^{a_k}$ for λ_k the distinct evals of α, so $a_1 + \cdots + a_k = n$. Let $m_\alpha(t) = (t-\lambda_1)^{c_1} \ldots (t-\lambda_k)^{c_k}$; $c_j \le \alpha_j \forall j$ since $m_\alpha \mid \chi_\alpha$ and $1 \le c_j$ since for each λ_j $\alpha(v) = \lambda v$ for some $v \ne 0 \in V$, so for p any poly $p(\alpha)(v) = p(\lambda)v$, $\vec{0} = m_\alpha(\alpha)(v) = m_\alpha(\lambda)v$ so λ a root of m_α.

T28

This is essentially an expansion of T15; let $\chi_\alpha = (\lambda_1 - t)^{a_1} \ldots (\lambda_k - t)^{a_k}$, then α diagable iff $p(\alpha) = 0$ where $p(t) = (t - \lambda_1) \ldots (t - \lambda_k)$.

Rk29

Exercise: If $\chi_A(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_0$ then $a_0 = \det A, a_{n-1} = (-1)^{n-1} trA$.

Jordan normal form

The full proof of this is the highlight of the IB GRM course; we work over $F = \mathbb{C}$. The JNF is bidiagonal; it has nonzero entries on the diagonal and possibly some 1s immediately above the diagonal. It is block diagonal; it has a set of square blocks $B_1, \ldots, B_k$ along the diagonal where $\lambda_1, \ldots, \lambda_k$ are the distinct evals, and 0s elsewhere. If we fix j and look at $B = B_j$ this is also block diagonal, made up of blocks $C_1, \ldots, C_g$ (where $g = g_{\lambda_j}$ as defined above);

each of the C_i is a jordan block $J_{S_i}(\lambda)$, the $S_i \times S_i$ block with entries λ on the diagonal, 1 immediately above it, and 0 elsewhere. We can arrange to have $S_1 \geq S_2 \geq \cdots \geq S_g = 1$.

T30

Every mat in $M_n(\mathbb{C})$ is conj to one in JNF, essentially unique (i.e. unique up to the order of the λ_j). The proof is not examinable in this course; see GRM, but an outline is as follows:

T31 Primary Decomposition T

Let $m_\alpha(t) = (t-\lambda_1)^{c_1} \dots (t-\lambda_k)^{c_k}$, then $V = V_1 \oplus \cdots \oplus V_k$ where $V_j = N((\alpha - \lambda_j \iota)^{c_j})$, generalized eigenspaces. We proove this similarly to 15; write $p_j(t) = (t-\lambda_1)^{c_1} \dots (t-\lambda_{j-1})^{c_{j-1}} (t-\lambda_{j+1})^{c_{j+1}} \dots (t-\lambda_k)^{c_k}$, then $p_1, \dots, p_k$ are coprime polys so by an analogue of Bezout's Thm (see N&S) $\exists$ polys $q_1, \dots, q_k$ s.t. $p_1 q_1 + \cdots + p_k q_k = 1$; let $h_j = p_j q_j$, then $\iota = h_1(\alpha) + \cdots + h_k(\alpha)$; $V_j = Im(h_j(\alpha))$ is in fact $N((\alpha - \lambda_j \iota)^{c_j})$, and each V_j is α-invariant, so we can split the matrix into B_j as required. Then we consider the restriction of α to V_j which is equivalent to the case $\chi_\alpha(t) = (\lambda - t)^n$. $m_\alpha(t) = (t-\lambda)^c$; let $v \in V$ w/ $(\alpha - \lambda\iota)^{c-1}(v) \neq 0$ (must exist by def of c), then $(\alpha - \lambda\iota)^{c-1}(v), (\alpha - \lambda\iota)^{c-2}(v), \dots, (\alpha - \lambda\iota)(v), v$ are lin ind [by applying $\alpha - \lambda\iota$]; let them respectively $= v_1, \dots, v_c$. Restricting α to the sp $W = \langle v_1, \dots, v_c \rangle$ we have the mat $\begin{pmatrix} \lambda & 1 & 0 & \dots \\ 0 & \lambda & 1 & \dots \\ 0 & 0 & \lambda & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$; $\alpha(v_1) = \lambda v_1$ as $(\alpha - \lambda\iota) v_1 = 0$, then $(\alpha - \lambda\iota) v_2 = v_1$ so $v_2 = \lambda v_2 + v_1$ and so on. The proof is completed fully in GRM but the remainder is relatively uninteresting; we take an α-invariant complement U to W and then have the matrix for B_j as $\begin{pmatrix} J_c(\lambda) & 0 \\ 0 & \star \end{pmatrix}$ and induct.

Discussion, "uniqueness"

For the case $n = s$, $J_s(\lambda)$ represents α. Observe $\chi_\alpha(t) = (\lambda - t)^s$ and $m_\alpha(t) = (t-\lambda)^s$, since $(J_s(\lambda) - \lambda I)^k$ is the matrix with 1s k above the diagonal and 0s elsewhere; each time we multiply by $(J_s(\lambda) - \lambda I)$ we shift the row of 1s up one. From the matrix we can clearly see $(\alpha - \lambda\iota)^k$ has nullity k for $k \leq s$, s for $k > s$ (since the max possible nullity is s). We can use this for the general case; the number of blocks with $\lambda_j = \lambda$ of size $\geq k$ is $n\left((\alpha - \lambda\iota)^k\right) - n\left((\alpha - \lambda\iota)^{k-1}\right)$; it follows that:

L32

The no. of blocks w/ $\lambda_j = \lambda$ of size k is $2n\left((\alpha - \lambda\iota)^k\right) - n\left((\alpha - \lambda\iota)^{k-1}\right) - n\left((\alpha - \lambda\iota)^{k+1}\right)$, so the JNF of α (assuming it exists) is determined by these dimensions of nullspaces, so unique in the sense described above.

For example, the JNFs for 2×2 mats are $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ $(J_1(\lambda_1) \oplus J_2(\lambda_2))$ for $\lambda_1 \neq \lambda_2$ w/ min poly $(t - \lambda_1)(t - \lambda_2)$, $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ $(J_1(\lambda) \oplus J_2(\lambda);\ (t - \lambda))$ and $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ $(J_2(\lambda);\ (t - \lambda)^2)$; the reader may wish to look at the $n = 3$ case where again we can distinguish by min polys; also consider the $n = 4$ case where we have λ with multiplicity 4; notice how fast the number of possible cases grows.

So e.g. if we know $m_\alpha(t) = (t - \lambda)^2$ in a 2D space we know $[\alpha]_B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ for some B; we can find B by taking v_2 (for the second col) $\in V \backslash N(\alpha - \lambda\iota)$ and $v_1 = (\alpha - \lambda\iota) v_2$.

JNF is very nice - given a mat A in it we can immediately see χ_A, m_A and for any ev λ a_λ, C_λ the size of the biggest λ-block, and g_λ the no. of λ-blocks.

5 Dual Sps, Dual maps

V is a fin dim vec sp over F in this sec unless otherwise specified. We def $V^\star = L(V, F)$ i.e. $\{\alpha : V \to F \text{ linear}\}$ the dual of V; this is a vec sp over F with elts these maps, linear functionals.

Let $B = e_1, \ldots, e_n$ some basis of V, then $B^\star = \epsilon_1, \ldots, \epsilon_n$ where ϵ_j is the linear extension of $\epsilon_j(e_k) = \delta_{jk}$ is the basis dual to B; lin ind as if $\left(\sum_j \lambda_j \epsilon_j\right)(e_k) = 0$ $\sum_j \lambda_j \delta_{jk} = 0$, span by $\alpha = \sum_j \alpha(e_j) \epsilon_j$; this implies $\dim V^\star = \dim V$.

C3

$\dim V^\star = \dim V$

It is sometimes useful to think of $V^\star$ as the sp of rows of length n over U; if $e_1, \ldots, e_n$ a basis of V and $\epsilon_1, \ldots, \epsilon_n$ the dual basis to it and $x \in V = \sum x_i e_i, \alpha \in V^\star = \sum a_i \epsilon_i$ then $\alpha(x) = \sum a_i x_i$ which we can see as the mat prod $(a_1, \ldots, a_n) \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$.

D4

If $U \subset V$ def U^0 the set of $\alpha \in V^\star$ such that $\alpha(u) = 0 \forall u \in U$, the anhilliator [sp] of U.

L5

If $U \subset V$, $U^0 \leq V^\star$; if $U \leq V$, $\dim V = \dim U + \dim U^0$ as take $e_1, \dots, e_k$ basis of U and extend to $B = e_1, \dots e_n$ basis of V, let $\epsilon_1, \dots, \epsilon_n$ the dual basis of $V^\star$, then $U^0 = \langle \epsilon_{k+1}, \dots, \epsilon_n \rangle$ as if $i > k$, $\epsilon_i (e_j) = 0 \forall j \leq k$ so $\epsilon_i \in U^0$ and if $\alpha \in U^0$ write $\alpha = \sum_{i=1}^n \lambda_i \epsilon_i$ and then for $i \leq k$ $\alpha (e_i) = 0$ so $\lambda_i = 0$ and $\alpha = \sum_{i=k+1}^n \lambda_i \epsilon_i$.

L6

Let W a vec sp over F and $\alpha \in L(V, W)$ then $\alpha^\star : W^\star \to V^\star$ given by $\epsilon \mapsto \epsilon \circ \alpha$ is linear; we call it the dual of α; exists since $\epsilon \circ \alpha$ is a lin map $V \to F$ so $\in V^\star$ and linear trivially.

Prop 7

Let $B = \{b_1, \dots, b_n\}, C$ bases of V, W respectively w/ respective dual bases $B^\star = \{\beta_1, \dots, \beta_n\}, C^\star$. For $\alpha \in L(V, W)$ $[\alpha^\star]_{C^\star B^\star} = [\alpha]_{BC}^T$; let $[\alpha]_{BC} = A = (a_{ij})$; $\alpha (b_j) = \sum_i a_{ij} c_i \forall j$ [c_j in my notes but this must be wrong], then $(\alpha^\star (\gamma_r))(b_s) = (\gamma_r \circ \alpha)(b_s) = \gamma_r \left(\sum_t a_{ts} c_t\right) = \sum_t a_{ts} \gamma_r (c_t) = \sum_t a_{ts} \delta_{rt} = a_{rs} = \sum_i a_{ri} \beta_i (b_s) \forall s$ so $\alpha^\star (\gamma_r) = \sum_i a_{ri} \beta_i \forall r$ and done.

C8

If $\dim V = \dim W$ $\det(\alpha) = \det(\alpha^\star), \chi_{\alpha^\star} = \chi_\alpha, m_{\alpha^\star} = m_\alpha$ (for any poly p over f, $p\left(A^T\right) = (p(A))^T$).

L9

$N(\alpha^\star) = (Im(\alpha))^0$ (so in particular $\alpha^\star$ inj iff α surj) as $\epsilon \in W^\star$ is $\in N(\alpha^\star)$ iff $\alpha^\star(\epsilon) = 0$ iff $\epsilon \circ \alpha = 0$ iff $\epsilon \in (Im(\alpha))^0$ and done.

Similarly, $Im(\alpha^\star) = (N(\alpha))^0$; for $\epsilon \in Im(\alpha^\star)$ $\epsilon = \alpha^\star(\phi)$ some $\phi \in w^\star$; for any $u \in N(\alpha)$ $\epsilon(u) = (\alpha^\star(\phi))(u) = \phi(\alpha(u)) = \phi\left(\vec{0}\right) = 0$ so $\epsilon \in (N(\alpha))^0$ and $Im(\alpha^\star) \supset (N(\alpha))^0$ and then equality by dimensions.

C10

$r(\alpha) = r(\alpha^\star)$ (so $r(A) = r\left(A^T\right)$, another proof of T2.29); $r(\alpha^\star) = \dim W^\star - n(\alpha^\star) = \dim W - \dim(Im(\alpha))^0 = \dim W - (\dim W - \dim Im(\alpha)) = r(\alpha)$.

For $v \in V$ let $\hat{v}(\epsilon) = \epsilon(v)$, the evaluation at v map; this is $\in V^{\star\star}$.

T11

$\hat{}: V \to V^{\star\star}$ as defined above is an isomorphism; note that this is a "natural" isomorphism without reference to bases. $\hat{}$ does map $V \to V^{\star\star}$ since $\hat{v} : V^\star \to F$ linear $\forall v \in V$, is trivially linear, injective by if $e \neq \vec{0} \in V$ let $e, e_2, \dots, e_n$ a basis of V and $\epsilon_1, \ldots, \epsilon_n$ the dual basis of $V^\star$, then $\hat{e}(\epsilon_1) = \epsilon_1(e) = 1$ so $\hat{e} \neq 0$, $\hat{}$ linear so inj; surj by dimensions so $\hat{}$ is an iso.

Rk12

If $\epsilon_1, \ldots, \epsilon_n$ a basis of $V^\star$ and $E_1, \ldots, E_n$ the basis of $V^{\star\star}$ dual to it $E_j = \hat{e_j}$ for unique $e_j \in V$; then $\epsilon_1, \ldots, \epsilon_n$ is the basis of $V^\star$ dual to $e_1, \ldots, e_n$.

L13

Let $U \leq V$, then $\hat{U} = U^{00}$; if we identify V with $V^{\star\star}$ by $\hat{}$, $U^{00} = U$; $U \leq U^{00}$ since $u \in U \Rightarrow \epsilon(u) = 0 \forall \epsilon \in U^0$ by def of U^0 so $\hat{u}(\epsilon) = 0 \forall \epsilon \in U^0$ by def of $\hat{}$ so $\hat{u} \in U^{00}$; equality by dimensions.

Rk14

For $T \leq V^\star$ we can def T^0 by $\{v \in V : \theta(v) = 0 \forall \theta \in T\}$.

L15

For $U_1, U_2 \leq V$, $(U_1 + U_2)^0 = U_1^0 \cap U_2^0$ (exercise), then applying ^{0} to this, $(U_1 \cap U_2)^0 = U_1^0 + U_2^0$.

Rk16

Let $V = P$ the set of all real polys; $P = \langle p_0, p_1, \ldots \rangle$ where $p_j(t) = t^j$; any $\epsilon \in P^\star$ can be written as $(\epsilon(p_0), \epsilon(p_1), \ldots) \in \mathbb{R}^n$ and all such sequences can be attained (see Exs3Q16) but $\mathbb{R}^{\mathbb{N}}$ has no countable generating set, and its dual will be even bigger, so cannot be iso to P. So these proofs really do depend on V being fin dim.

Rk17

We have a mapping $V^\star \times V \to F$ by $(\epsilon, v) \mapsto \epsilon(v)$; this is a bilinear func on $V^\star \times V$ (see later); we write it as $\langle \epsilon \mid v \rangle$ as we could equally well use $\hat{v}(\epsilon)$ so this is symmetric; we have $\langle \alpha^\star(\epsilon) \mid v \rangle = \langle \epsilon \mid \alpha(v) \rangle$ ($\forall \alpha$ as above)

6 Bilinear Forms

In this section V, W vec sps over F, fin dim unless otherwise specified

Def

The func $\psi : V \times W \to F$ is a bilinear func if it is linear in each coordinate, i.e. $\psi(v, w)$ is linear in v $\forall$ fixed $w \in W$ and vv; here we usually take $V = W$ in which case we say ψ is a bilinear form (on V). For example, the real inner or scalar product on $\mathbb{R}^n \times \mathbb{R}^n$, or more generally for $V = F^n$ and A fixed $\in M_n(F)$, $\psi(u, v) = u^T A v$ is a bilinear form on V.

D3

Let $\dim V = n$ and $B = v_1, \ldots, v_n$ a basis of V, then the mat of the bilinear form ψ on V wrt B is $[\psi]_B = (\psi(v_i, v_j))$ as an $n \times n$ matrix.

L4

$\psi(u, v) = [u]_B^T [\psi]_B [v]_B \, \forall u, v \in V$; furthermore $[\psi]_B$ is the only mat for which this holds; $\psi(a, b) = \psi\left(\sum a_i v_i, \sum b_j v_j\right) = \sum a_i b_j \psi(v_i, v_j) = (a_1, \ldots, a_n) [\psi]_B \begin{pmatrix} b_1 \\ \ldots \\ b_n \end{pmatrix}$ and done; if $\psi(u, v) = [u]_B^T A [v]_B \, \forall u, v \in V$ take $u = v_i, v = v_j$ and $A = [\psi]_B$.

Change of basis

For $B' = v'_1, \ldots, v'_n$ also a basis of V and P the change of basis mat from B to B' $v_j = \sum p_{ij} v_i$ and $[v]_B = P [v]_{B'} \, \forall v \in V$.

T5

$[\psi]_{B'} = P^T [\psi]_B P$ (note P^T rather than P^{-1}) as $\psi(u, v) = [u]_B^T [\psi]_B [v]_B = \left(P [u]_{B'}\right)^T [\psi]_B P [v]_{B'} = [u]_{B'}^T P^T [\psi]_B P [v]_{B'}$ and done.

D6

Square real $n \times n$ mats A, B have A congruent to B if $B = P^T A P$.

L7

This is an equiv rel on $M_n(\mathbb{R})$; A cong A by $P = I$, if A cong B by P then B cong A by P^{-1}, and if also B cong C by Q then A cong C by PQ

D8

The rank of a bilinear form $r(\psi)$ is $r([\psi]_B)$ for any basis B; this is well defd.

D9

A real bilinear form ψ on V is symmetric if $\psi(u, v) = \psi(v, u) \, \forall u, v \in V, \lambda \in F$; note it is equiv that $[\psi]_B$ is diagonal; To be able to represent ψ by a diagonal mat ψ must be symmetric as if $P^T A P = D$ diagonal, $D = D^T = P^T A^T P$ so $A = A^T$ (since P invertible).

D10

For V a real vec sp $Q : V \to \mathbb{R}$ is a quadratic form if $Q(\lambda v) = \lambda^2 Q(v), \exists$ a real symmetric bilinear form ψ st $Q(u+v) = Q(u) + Q(v) + 2\psi(u,v)$; note we can find ψ given Q by $\psi(u,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v))$ or $\frac{1}{4}(Q(u+v) - Q(u-v))$, and for any ψ we have a corresponding Q by $Q(v) = \psi(v,v)$;

T11

Any real symmetric bilinear form (or equivalently any real quadratic form, as is the case for many of the following results) can be represented by a diagonal mat; moreover this can be taken to be $\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$ for some $p, q \in \mathbb{N}_0$; given a real symmetric bilinear form ψ on V $\exists$ a basis B of V s.t. if $[v]_B = \begin{pmatrix} X_1 \\ \dots \\ X_n \end{pmatrix}$, $Q(V) = X_1^2 + \cdots + X_p^2 - X_{p+1}^2 - \cdots - X_{p+q}^2$; we induct on $\dim V$; we can assume ψ is nonzero, otherwise we are done; then $\exists v \in V$ with $Q(v) \neq 0$; then consider $W = \{w \in V : \psi(v,w) = 0\}$; $W \subsetneqq V$ since $v \notin W$ and it suffices to show $V = \langle v \rangle \oplus W$; if $u \in V$ we can write $u = \lambda v + (u - \lambda v)$; choose $\lambda \in \mathbb{R}$ so $u - \lambda v \in W$ by $\lambda = \frac{\psi(u,v)}{\psi(v,v)}$ so $V = \langle v \rangle + W$, and $\langle v \rangle \hat{W} = \vec{0}$ since if $\psi(\lambda v, v) = 0$ then $\lambda Q(v,v) = 0$ so $\lambda = 0$; now the restriction of ψ to $W \times W$ is a real symmetric bilinear form so we can induct (the base case is trivial [or so claims the lecturer]); we have a basis $B' = v_2, \dots, v_n$ in which it is diagonal and then ψ is diagonal wrt $B = v, v_2, \dots, v_n$.

Let $[\psi]_B = \begin{pmatrix} d_1 & & \\ & \dots & \\ & & d_n \end{pmatrix}$; reorder B if necessary so that the first p of the d_i are $+ve$, the next q $-ve$ and the rest 0; then normalize B by $v_i \to \frac{v_i}{\sqrt{Q(v_i)}}$ for $1 \leq i \leq p$, $\frac{v_i}{\sqrt{-Q(v_i)}}$ for $p+1 \leq i \leq p+q$; then the mat of ψ wrt this new B is as required.

D12

As per above, the rank $r(\psi) = p + q$; signature $s(\psi) = p - q$, and these are basis-invariant:

T13 Sylvester's Law of Inertia

If a real symmetric [bilinear] form ψ is represented by $\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} I_{p'} & 0 & 0 \\ 0 & -I_{q'} & 0 \\ 0 & 0 & 0 \end{pmatrix}$ wrt bases B, B' then $p = p', q = q'$; first def Q on a real vec sp V w/ $U \leq V$ is

$+ve$ definite on U if $Q(u) > 0 \forall u \neq \vec{0} \in U$, $+ve$ semidefinite for $\geq$ rather than $>$, similarly $-ve$ definite and semidefinite; if we say Q $+ve$ definite without additional qualification we mean Q is $+ve$ definite on V and similarly. Now, we claim p is the largest dim of a subsp on which ψ is $+ve$ definite (and sim q for $-ve$ definite); we sometimes define p, q by these; $B = v_1, \ldots, v_p, v_{p+1}, \ldots, v_{p+q}, \ldots, v_n$; let $P = \langle v_1, \ldots, v_p \rangle, U = \langle v_{p+1}, \ldots, v_n \rangle$; ψ is $+ve$ definite on P, and if ψ $+ve$ definite on some P', $P' \cap U = \left\{ \vec{0} \right\}$ since ψ $-ve$ semidefinite on U so $\dim P' \leq \dim V - \dim U$, $p' + n - p \leq n$ so $p' \leq p$ and the claim holds; sim for q. Of course this is equivalently true for a real quadratic form over $\mathbb{R}$.

Rk14

ψ determines p but not P; there are generally many possible such spaces, sim for q; note that rank and signature together determine p, q. $K = \langle v_{p+q+1}, \ldots, v_n \rangle$ is determined by ψ; it is the kernel or radical of the form: $K = \{v \in V : \psi(v, u) = 0 \forall u \in V\}$; we call it $V^{\perp}$.

Def

ψ is non-singular if $K = \left\{ \vec{0} \right\}$ or equivalently $r(\psi) = \dim V$; note we may still have $U \subset V$ with $\psi(u, v) = 0 \forall u, v \in U$.

Rk15

$\exists$ subsp T of dim $\min\{p, q\} + n - (p + q)$ s.t. $\psi = 0$ on T; this includes K but is generaly much larger. $\min\{p, q\} + n - (p + q)$ is the largest possible dim of such a sp; say wlog $q \leq p$ and take $T = \langle v_1 + v_{p+1}, \ldots, v_q + v_{p+q}, v_{p+q+1}, \ldots, v_n \rangle$ (note $T \cap P = \{0\} = T \cap Q$).

e.g. if ψ is non-singular with $n = 2m$ and $\exists$ a subsp of dim m on which ψ is 0 then $p = m = q$ so $s(\psi) = p - q = 0$.

16 Worked Example

$V = \mathbb{R}^3, Q(\vec{x}) = x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3$; the mat of Q wrt the standard basis is $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}$ (this can be found by $\psi(u, v) = \frac{1}{2}(Q(u + v) - Q(u) - Q(v))$). The first method to diagonalise is to gather all ocurrences of x_1 together as $Q(\vec{x}) = (x_1 + x_2 + x_3)^2 + x_3^2 - 4x_2x_3$, then all x_3 [since this is easier] by $Q(\vec{x}) = (x_1 + x_2 + x_3)^2 + (x_3 - 2x_2)^2 - (2x_2)^2$ (in fact this offers another way to proove T11) so we know $[Q]_B = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$ wrt some B; $r(\psi) = 3, s(\psi) = 1$. Then to find a suitable trans mat P we have

$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ so $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & 0 \end{pmatrix}^{-1}$. For the second method we apply eltary col ops followed by the corresponding eltary row op $A \to E^T AE$ e.g. $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}$; we want col 2 $\to$ col 2 - col 1 so $E_1 = \begin{pmatrix} 1 & -1 & \\ & 1 & \\ & & 1 \end{pmatrix}$, then $AE_1 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & -2 & 2 \end{pmatrix}$, $E_1^T AE_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -2 \\ 1 & -2 & 2 \end{pmatrix}$; we similarly use col 3 $\to$ col 3 - col 1 and so on; we build up P as we go along by $E_1 E_2 \dots$. For the third method we can use the same method as the pf of T11.

Finally if we just want to find r, s it is sometimes easier to work with χ_A since we shall later see s is the no. of $+ve$ evals of A $-$ the no. of $-ve$ evals of A.

Now we work over $F = \mathbb{C}$; for ψ bilinear and symmetric on V over $\mathbb{C}$ as in T11 we have a basis B s.t. $[\psi]_B = \begin{pmatrix} d_1 & & & & & \\ & \dots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \dots & \\ & & & & & 0 \end{pmatrix}$ w/ $d_i \neq 0 \in \mathbb{C} \forall i$; now replace v_i by $\frac{v_i}{\sqrt{v_i}} \forall 1 \leq i \leq r$ and then ψ has mat $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ wrt this new basis, so:

L17

Any cplx symmetric mat A satisfies $P^T AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ for some invertible P for unique r (actually $r(A)$); this is usially not quite what we want. Rather than symmetric cplx mats we need to study Hermitian mats; a mat A is Hermitian if $A = \overline{A^T}$ (complex conjugation).

D18

For V a cplx vec sp a Hermitian form on V is a func $\psi : V \times V \to \mathbb{C}$ s.t. $\forall v \in V, u \mapsto \psi(u, v)$ is linear (note this is the ohter way around from in QM) and $\psi(u, v) = \overline{\psi(v, u)}$. Note that such a ψ is not a bilinear form on V, rather it is sesquilinear [sp?]: $\psi(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 \psi(u_1, v) + \lambda_2 \psi(u_2, v), \psi(u, \lambda_1 v_1 + \lambda_2 v_2) = \overline{\lambda_1} \psi(u, v_1) + \overline{\lambda_2} \psi(u, v_2)$; an example of such a form is the cplx inner prod.

Rk21

For V a cplx vec sp and ψ a Herm form on V, can def $Q : V \to \mathbb{C}$ (in fact Q is real-valued) by $Q(v) = \psi(v, v)$; we have $Q(\lambda v) = |\lambda|^2 Q(v)$; given Q we can recover ψ similarly to before; $\psi(u, v) = \frac{1}{4}(Q(u+v) - Q(u-v) + iQ(u+iv) - iQ(u-iv))$. If $B = v_1, \ldots, v_n$ is a basis of V the mat of ψ wrt B is $[\psi]_B = (\psi(v_i, v_j))$. Let this be A, then $A = \overline{A^T}$ i.e. this is a Herm mat. $\psi(u, v) = [u]_B^T [\psi]_B \overline{[v]_B}$.

Finally , a change of basis maps $[\psi]_B \to P^T A \overline{P}$ where P is the (invertible) change of basis mat.

T26

If ψ is a Herm form on the cplx vec sp V $\exists$ a basis B of V w/ $[\psi]_B = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}$, and p, q are determined by ψ. The proof is mostly as for the reals; as an outline if $\psi \equiv 0$ we are done, otherwise take $v \neq \vec{0} \in V$ w/ $\psi(v, v) \neq 0$, then def $W = \{w \in V : \psi(v, w) = 0\}$ and $V = \langle v \rangle \oplus W$ since if $u \in V$ $u = \lambda v + (u - \lambda v)$ with $\lambda = \frac{\psi(u,v)}{\psi(v,v)}$ so $\psi(v, u - \lambda v) = 0$; we then inductively find $v_2, \ldots, v_n$ a basis of W wrt which $\psi|_W$ is diagonal, then take $B = v_1, v_2, \ldots, v_n$ and $[\psi]_B$ is diagonal; the top row is 0s other than the top left so since the mat is Herm the left collumn is also all 0 below the top. Then we reorder the basis so the first p entries are $+ve$, the next q $-ve$ and the rest 0, then replace v_j by $\frac{1}{\sqrt{|Q(v_j)|}} v_j$ for j from 1 to $p+q$. That p, q are determined is by exactly the same pf as in 13; p is the maximal dim of a subsp on which ψ is $+ve$ definite etc.

Returning to V a real vec sp, there is another important class of real bilinear forms:

D27

The bilinear form ψ on the real vec sp V is skewsymmetric or symplectic or alternating if $\psi(v, u) = -\psi(u, v) \, \forall u, v \in V$; note this means $\psi(v, v) = 0 \forall v \in V$. If $A = [\psi]_B$ for some basis B of V then $A^T = -A$; A is skewsymmetric.

Rk28

Any real square mat A can be written as $A = \frac{1}{2}\left(A + A^T\right) + \frac{1}{2}\left(A - A^T\right)$ a sum of symmetric and antisymmetric parts.

T29

If ψ is a skewsymmetric bilinear form on a real vec sp V then m\exists a basis $v_1, w_1, v_2, w_2, \ldots, v_m, w_m, v_{2m+1}, v_{2m+2}, \ldots, v_n$ wrt which ψ has mat

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

note that this means the rank of any skewsymmetric mat is even. Also note we can rearrange the basis as $v_1, \ldots, v_m, w_1, \ldots, w_m, v_{2m+1}, \ldots, v_n$ and the mat becomes $\begin{pmatrix} 0 & I_m & \\ -I_m & 0 & \\ & & 0 \end{pmatrix}$; for the proof we induct on $\dim V = n$; if $\psi \equiv 0$ we are done, otherwise $\exists$ vecs v_1, w_1 w/ $\psi(v_1, w_1) \neq 0$ and by scaling w_1 we can have this $= 1$, then $\psi(w_1, v_1) = -1$. Let $U = \langle v_1, w_1 \rangle$ and $W = U^\perp = \{v \in V : \psi(v_1, v) = 0 = \psi(w_1, v)\}$; then $V = U \oplus W$ as given $v \in V$ let $a = \psi(v, w_1), b = \psi(v_1, v)$ then $v = av_1 + bv_2 + (v - av_1 - bw_1)$ with the first two terms in U and the last in W, and $U \cap W = \left\{\vec{0}\right\}$ as if $av_1 + bw_1 \in W$ then $\psi(v_1, aw_1 + bv_1) = 0 \Rightarrow a = 0, \psi(w_1, aw_1 + bv_1) \Rightarrow b = 0$. Now we continue with $\psi|_W$ and the claim holds by induction.

Some extra remarks on general bilinear forms on $U \times V$ where U, V are over the same field F; $\psi : U \times V \to F$ linear in each coordinate. [The lecturer used to be making a distinction between forms and functions, and made a big fuss about this, but appears to have now abandoned this. Or just be incompetent. Or both] Examples are $U = V^\star, \psi$ given by $(\alpha, v) \mapsto \alpha(v)$ or for V over $F = \mathbb{C}$ define $\overline{V}$ to have the same elts as V and the same addition but with scalar prod $\lambda \bar{\cdot} v = \overline{\lambda} \cdot v$; the sesquilinear forms on V are precisely the bilinear funcs $V \times \overline{V} \to \mathbb{C}$.

Cor 30

The rk of any skewsymmetric mat is even.

For U, V over F and $\psi : U \times V \to F$ bilinear. we have maps $\psi_L : U \to V^\star$ by $u \mapsto \psi_L(u)$ given by $v \mapsto \psi(u, v)$, sim. $\psi_R : V \to U^\star$. ψ is non-singular if both the left kernel $\ker \psi_L$ and right kernel $\ker \psi_R$ are $\{0\}$.

L31

If ψ is non-singular on $U \times V$ then $\dim U = \dim V$ as $\ker \psi_L = \{0\} \Rightarrow \dim U \leq \dim V^\star = \dim V$ and similarly $\dim V \leq \dim U$.

L32 (Exercise)

If $\dim U = \dim V$ then $\ker \psi_L = \{0\} \Leftrightarrow \ker \psi_R = \{0\}$; in fact if we assume $\ker \psi_L = \{0\}$ let $u_1, \ldots, u_n$ be a basis of U, then $\psi_L(u_1), \ldots, \psi_L(u_n)$ is a basis of $V^\star$; let $v_1, \ldots, v_n$ the basis of V dual to it and observe $\psi(u_i, v_j) = \delta_{ij}$, so we have bases of U, V which are "dual wrt ψ".

33

Let ψ a non-singular bilinear form on V, then $\psi_L : V \to V^\star$ is an isomorphism.

34

For ψ a non-singular bilinear form V and $W \leq V$, then $W^\perp$ (the right [perp, I assume - lol saxl's accent] of W) is $\{v \in V : \psi(w, v) = 0 \forall w \in W\}$. We clearly have $W^\perp \leq V$; we claim $\dim V = \dim W + \dim W^\perp$ which is true since $W^\perp = (\psi_L(W))^0$, as $v \in W^\perp \Leftrightarrow \psi(w, v) = 0 \forall w \in W \Leftrightarrow \psi_L(w)(v) = 0 \forall w \in W \Leftrightarrow v \in (\psi_L(W))^0$, so then $\dim W + \dim W^\perp = \dim W + \dim(\psi_L(W))^0 = \dim W + \dim V - \dim \psi_L(W) = \dim V$. ($\psi$ is non-singular so $\dim \psi_L(W) = \dim W$)

7 Inner Product Sps

D1

For V a real/cplx vecsp an inner prod on V is a +ve definite symmetric bilinear/Herm form on V; as notation we write $\langle v, w \rangle$ for the value of the inner prod on (v, w). If V is a real/cplx inner prod sp (i.e. a sp w/ an inner prod) it is a Euclidean/unitary sp. (i.e. a real one is Euclidean, a cplx one is unitary, and similarly), e.g. dot products, or (exercise) $V = C[0, 1]$ the spare of cnts real- or cplx-vald funcs on $[0, 1]$ w/ $\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt$.

D2

The length $\|v\|$ of $v \in V$ is $+\sqrt{\langle v, v \rangle}$; note $\langle v, v \rangle \geq 0$ w/ equality iff $v = \vec{0}$.

L3 The schwartz ineq

$|\langle v, w \rangle| \leq \|v\| \|w\| \forall v, w \in V$; if $v = 0$ trivial, otherwise for the real case $0 \leq \|tv - w\|^2 = t^2 \|v\|^2 - 2t \langle v, w \rangle + \|w\|^2 \forall t \in \mathbb{R}$ (here we could use the discriminant of this quadratic in t, but we want to use a similar proof for both cases); put $t = \frac{\langle v, w \rangle}{\|v\|^2}$ and then $0 \leq -\frac{\langle v, w \rangle^2}{\|v\|^2} + \|w\|^2$ so $|\langle v, w \rangle| \leq \|v\| \|w\|$, and for the cplx case $0 \leq \|tv - w\|^2 = t\bar{t} \|v\|^2 - (t + \bar{t}) \langle v, w \rangle + \|w\|^2 \forall t \in \mathbb{C}$; put $t = \frac{\overline{\langle v, w \rangle}}{\|v\|^2}$ then $0 \leq -\frac{|\langle v, w \rangle|^2}{\|v\|^2} + \|w\|^2$ and done as before.

D4

In the Euclidean case, if $v \neq \vec{0} \neq w$ the angle θ between v, w is given by $\cos\theta = \frac{\langle v,w \rangle}{\|v\|\|w\|}$, taking $\theta \in [0, \pi]$.

L5 Triangle ineq

$\|v + w\| \leq \|v\| + \|w\|$ as $\|v + w\|^2 = \|v\|^2 + \left(\langle v, w \rangle + \overline{\langle v, w \rangle} \right) + \|w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$.

D6

A set $e_1, \ldots, e_k$ of vecs $\in V$ is orthogonal if $\langle e_i, e_j \rangle = 0 \forall i \neq j$ and orthonormal if $\langle e_i, e_j \rangle = \delta_{ij} \forall i, j$.

L7

If $e_1, \ldots, e_j$ are orthog nonzero vecs they are lin ind; in fact $v = \sum \lambda_j e_j \Rightarrow \lambda_j = \frac{\langle v, e_j \rangle}{\langle e_j, e_j \rangle}$.

By 6.11,6.26 $\exists$ ON bases; there is a procedure for "making" them:

T8 The Gram Schmidt Orthogonalization Process

Let V an inner prod sp (always fin dim from now on); let $v_1, \ldots, v_n$ a basis of V. There is an ON basis $e_1, \ldots, e_n$ s.t. *span* $\langle e_1, \ldots, e_k \rangle =$ *span* $\langle v_1, \ldots, v_k \rangle \forall 1 \leq k \leq n$; let $e_1 = \frac{v_1}{\|v_1\|}$ and induct; if we have found $e_1, \ldots, e_k$ take $e'_{k+1} = v_{k+1} - \sum_{j=1}^{k} \lambda_j e_j$ w/ λ_j chosen so that $\left\langle e_j, e'_{k+1} \right\rangle = 0 \forall 1 \leq j \leq k$ by $\lambda_j = \langle e_j, v_{k+1} \rangle$. Then $e'_{k+1} \neq 0$ since $v_1, \ldots, v_{k+1}$ indep; put $e_{k+1} = \frac{e'_{k+1}}{\|e'_{k+1}\|}$ and done.

C9

In a fin dim inprosp [shorthand for inner product sp] any ON set of vecs can be extended to an ON basis; if $e_1, \ldots, e_k$ ON they are lin ind, extend to a basis $e_1, \ldots, e_k, v_{k+1}, \ldots, v_n$ and apply Gram Schmidt - first k vecs are unchanged since ON already.

D10

Let V and inprosp; if$W \leq V$ write $W^{\perp} = \{v \in V : v \perp w \forall w \in W\}$, where $v \perp w$ means $\langle v, w \rangle = 0$ or equivalently $\langle w, v \rangle = 0$. This is the orthogonal complement of W in V; it is clearly unique, an:

T11

If V a find dim inprosp, $W \leq V$ then $W^{\perp} \leq V$ and $V = W \oplus W^{\perp}$; let $e_1, \ldots, e_k$ an ON basis of W, extend this to an ON basis $e_1, \ldots, e_n$ of V; observe $e_{k+1}, \ldots, e_k \in W^{\perp}$ and $W^{\perp} = \langle e_{k+1}, \ldots, e_n \rangle$; if $v \in V$ we can write $v = \sum_{j=1}^{n} \lambda_j e_j = \sum_{j=1}^{k} \lambda_j e_j + \sum_{j=k+1}^{n} \lambda_j e_j$ so $V = W + W^{\perp}$; observe $W \cap W^{\perp} = \{0\}$ since if $v \in W \cap W^{\perp}$ $\langle v, v \rangle = 0$. From now on take $W \leq V$

D12

Any $v \in V$ can be written uniquely as $v = w + w'$ w/ $w \in W, w' \in W^{\perp}$. Def $\pi : V \to W$ by $v \mapsto$ this w; this is linear and surj, called the orthogonal projection of V onto W. It is a projection since $\pi^2 = \pi$. Also observe $\ker \pi = W^{\perp}$ and $\pi' = \iota - \pi$ is the orthog proj of V onto $W^{\perp}$.

L13

If $e_1, \ldots, e_k$ an ON [merely orthogonal in lectures, but that must be wrong] basis of W then π satisfies $\pi(v) = \sum_{j=1}^{k} \langle v, e_j \rangle e_j \forall v \in V$, as if $v = \sum_{j=1}^{n} \lambda_j e_j$ (extending to an ON basis of V) then $\lambda_j = \langle v, e_j \rangle$ and $\pi(v) = \sum_{j=1}^{k} \pi_j e_j$ since $v = \pi(v) \in W^{\perp}, = \sum_{j=1}^{k} \langle v, e_j \rangle e_j$. Note $\pi(v)$ is the point of W nearest to v; $d(v, \pi(v))$ (or $\|v - \pi(v)\|$) $\leq d(v, w) \forall w \in W$.

P14

Any real nonsingular (note therefore square) mat A can be written $A = RT$ where R is an orthog mat (i.e. $R^{-1} = R^T$) and T is upper triangular; sim for A cplx but then R is unitary ($R^{-1} = \overline{R^T}$). Work in $V = \mathbb{R}^n$ where A is $n \times n$, w/ standard dot prod. Let $v_1, \ldots, v_n$ the cols of A; this is a basis of V since A is nonsingular. Apply Gram-Schmidt; let $e_1, \ldots, e_n$ be the ON basis this obtained. Let R be the mat w/ cols $e_1, \ldots, e_n$, then $R^T R = I$ since the e_j are ON. Write $v_k = \sum_{j=1}^{n} t_{jk} e_j$ and let $T = (t_{ij})$; then T is upper triangular since $v_k \in span \langle e_1, \ldots e_k \rangle \forall k$ and $A = RT$ since $A^{(k)} = v_k = \sum_{j=1}^{n} t_{jk} R^{(j)}$.

Endomorphisms of inprosps

For V an inprosp and $\alpha : V \to V$ linear:

P15 (Important)

For V fin dim $\exists!$ endomorphism $\alpha^{\star}$ of V s.t. $\langle \alpha v, w \rangle = \langle v, \alpha^{\star} w \rangle \forall v, w \in V$; moreover for B an ON basis of V $[\alpha^{\star}]_B = \overline{[\alpha]_B^T}$. This is the adjoint of α; note that this is not the same as the $\alpha^{\star} : V^{\star} \to V^{\star}$ defined above (even though this is sometimes called the classical adjoint); the notation is standard in both cases. Let $B = e_1, \ldots, e_n$ an ON basis of V, $A = [\alpha]_B$, and let $\alpha^\star$ be the endomorphism of V given by $[\alpha^\star]_B = \overline{A^T} = C$, then $\forall 1 \leq i,j \leq n$, $\langle \alpha(e_i), e_j \rangle = \langle \sum_{k=1}^n a_{ki} e_i, e_j \rangle = \sum_{k=1}^n a_{ki} \delta_{kj} = a_{ji}$; $\langle e_i, \alpha^\star(e_j) \rangle = \langle e_i, \sum_{k=1}^n c_{kj} e_k \rangle = \sum_{k=1}^n \overline{c_{kj}} \delta_{ik} = \overline{c_{ij}}$, so by linearity $\langle \alpha(v), w \rangle = \langle v, \alpha^\star(w) \rangle \, \forall v, w$; uniqueness by the same proof in reverse: this property $\Rightarrow [\alpha^\star]_B = \overline{[\alpha]_B^T}$.

Rk16

For $F = \mathbb{C}$ put $\psi(v,w) = \langle v, w \rangle$, then $\psi_R(w) \in V^\star \forall w$, each given by $v \mapsto \psi(v,w)$; ψ_R is a map $V \to V^\star$. Then the map $V \to V^\star \to V^\star \to V$ given by $\psi_R^{-1} \circ \alpha^\star \circ \psi_R$ for $\alpha^\star$ the dual map of α is the adjoint map of α on V; if we identify $V, V^\star$ under ψ_R then the adjoint and the dual of α are the same thing, since $\left\langle v, \psi_R^{-1} \alpha^\star \psi_R w \right\rangle = \left(\psi_R\left(\psi_R^{-1}\left(\alpha^\star\left(\psi_R(w)\right)\right)\right)\right)(v) = \left(\alpha^\star\left(\psi_R(w)\right)\right)(v) = \left(\psi_R(w)\right)(\alpha(v)) = \langle \alpha(v), w \rangle \, \forall v, w \in V$; if we try and do the same with a cplx inprosp we get an identification of $\overline{V}$ with $V^\star$.

L17

For adjoint maps, $(\alpha + \beta)^\star = \alpha^\star + \beta^\star, (\lambda\alpha)^\star = \overline{\lambda}\alpha^\star, \alpha^{\star\star} = \alpha, \iota^\star = \iota$ either directly from the matricies or the direct proofs are trivial, e.g. $\langle v, \alpha^{\star\star}(w) \rangle = \langle \alpha^\star(v), w \rangle = \overline{\langle w, \alpha^\star(v) \rangle} = \overline{\langle \alpha(w), v \rangle} = \langle v, \alpha(w) \rangle \, \forall v, w \in V$ so $\langle v, (\alpha - \alpha^{\star\star}) w \rangle = 0 \forall v, w \in V$ i.e. $\alpha = \alpha^{\star\star}$.

D18

For V fin dim inprosp and $\alpha \in L(V)$ we define α is:

- Self-adjoint if $\alpha = \alpha^\star$; equivalently $\langle \alpha(v), w \rangle = \langle v, \alpha(w) \rangle \, \forall v, w \in V$; for V real α is symmetric, for V cplx α is Hermitian
- An isometry if $\alpha^\star = \alpha^{-1}$ or equivalently $\langle \alpha(v), \alpha(w) \rangle = \langle v, w \rangle \, \forall v, w \in V$; for V real α is orthogonal, for V cplx α is unitary
- Normal if $\alpha\alpha^\star = \alpha^\star\alpha$.

For matricies, a real matrix A is symmetric if $A^T = A$, orthogonal if $A^T = A^{-1}$ and a cplx mat A is Hermitian if $\overline{A^T} = A$ and unitary if $\overline{A^T} = A^{-1}$.

L19

If $\alpha \in L(V)$ for V a fin dim inprosp and B an ON basis therof, α is symmetric/hermitian/orthogonal/unitary iff $[\alpha]_B$ is.

L20

Let V a cplx inprosp, $\alpha \in L(V)$ Hermitian (unitary), then the evals of α are real (lie on the unit circle in $\mathbb{C}$ and evecs corresponding to distinct evals are orthogonal; if $\alpha(v) = \lambda v$ w/ $v \neq \vec{0}$ then $\lambda \langle v, v \rangle = \langle \alpha(v), v \rangle = \langle v, \alpha^\star(v) \rangle$ which

$= \langle v, \alpha(v)\rangle = \overline{\lambda}\langle v, v\rangle$; since $v \neq \vec{0}$ $\langle v, v\rangle \neq 0$ so this means $\lambda = \overline{\lambda}$ i.e. λ real $(= \langle v, \alpha^{-1}(v)\rangle = \overline{\lambda^{-1}}\langle v, v\rangle$ so $\lambda = \overline{\lambda^{-1}}$ so $|\lambda|^2 = 1)$. If $\alpha(v_i) = \lambda_i v_i$ for $i = 1, 2$ w/ $\lambda_1 \neq \lambda_2$ then $\lambda_1 \langle v_1, v_2\rangle = \langle \alpha(v_1), v_2\rangle = \langle v_1, \alpha^\star(v_2)\rangle = \langle v_1, \alpha(v_2)\rangle = \overline{\lambda_2}\langle v_1, v_2\rangle = \lambda_2 \langle v_1, v_2\rangle$ so $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$ and $\langle v_1, v_2\rangle = 0$ (the pf for unitary α is similar).

Main T: 21

Let V a cplx findim inprosp, α a Hermitian (unitary) endomorphism of V, then $\exists$ an ON basis of V consisting of evecs of α, i.e. $[\alpha]_B$ is diagonal wrt some ON B: since V is cplx α has an eval λ; let $\alpha(e) = \lambda e$ with $\|e\| = 1$ (which we can do by scaling since $e \neq \vec{0}$); let $W = \langle e\rangle^\perp$, then $V = \langle e\rangle \oplus W$ by T11 (or an easy direct pf) and $\alpha(W) = W$; W is α-invariant: if $v \in W$ then $\langle \alpha(v), e\rangle = \langle v, \alpha^\star(e)\rangle = \langle v, \alpha(e)\rangle = \overline{\lambda}\langle v, e\rangle = 0$ $(= \overline{\lambda^{-1}}\langle v, e\rangle = 0$ for unitary), so $\alpha(v) \in W$. Now $\alpha|_W$ is Hermitian (unitary) so by induction $\exists$ an ON basis $e_2, \ldots, e_n$ of W consisting of evecs of α and then $\{e, e_2, \ldots, e_n\}$ is an ON basis of V of evecs of α.

L22

Let V a real findim inprosp, $\alpha \in L(V)$ a symmetric endomorphism therov, then α has real evals and evecs corresponding to distinct evals are orthog; for B an ON basis of V $[\alpha]_B$ is a real symmetric mat so Hermitian so by 20 the evals of α are real, and we have orthogonality by the same pf as in 20.

Main T: 23

Note that this T does not in general work for orthog endomorphisms, only symmetric ones; let V a real findim inprosp, α a symmetric endomorphism therof, then $\exists$ an ON basis (of V) of evecs of α: by 22 α has a real eval so let e a corresponding evec of length 1, let $W = \langle e\rangle^\perp$ and continue as in T21

A common generalisation, which should be considered as an exercise: for V over $\mathbb{C}$ and $\alpha \in L(V)$ normal (i.e. $\alpha\alpha^\star = \alpha^\star\alpha$), $\exists$ an ON basis of evecs.

Rk24

L22 and hence T23 do not hold for orthog endomorphisms of real inprosps e.g. $n = 2$, α a rotation has in general no real evals; however, see Exs4Q14: for V a real inprosp and $\alpha \in L(V)$ orthogonal $\exists$ an ON basis B st $[\alpha]_B =$

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \Box & & \\ & & & & & & & \ddots & \\ & & & & & & & & \Box \end{pmatrix}$$ where the $\Box$ are $2{\times}2$ blocks $\begin{pmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{pmatrix}$;

as an outline of the pf if α has a real eval λ then $\lambda = \pm 1$ as before (λ must be on the unit circle as per L20). Assume all irreducible factors of the min poly m of α (i.e. α has no real evals), then let $m_\alpha(x) = \left(x^2 + ax + b\right) q(x)$; $q(\alpha) \neq 0$; let $v \in Im\,(q(\alpha))$, then $\left(\alpha^2 + a\alpha + b\right)(v) = \vec{0}$; let $W = [span]\,\langle v, \alpha(v)\rangle$, then $V = W \oplus W^\perp$; both W and $W^\perp$ are α-invariant, and we induct.

Rk 24A

Let $A \in M_n(\mathbb{R})$ $(M_n(\mathbb{C}))$ symmetric (Herm); regard it as an endomorphism of $\mathbb{R}^n$ $(\mathbb{C}^n)$ w/ standard inner prod: $v \mapsto Av$. $\exists$ an ON basis $v_1, \ldots, v_n$ of evecs by the above. Then $P = (v_1 \ldots v_n)$ is orthogonal (unitary) and $AP = Pd$ w/ D diagonal $= \begin{pmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_n \end{pmatrix}$; then $P^{-1}AP = D = P^TAP$ $(\overline{P^T}AP)$. Note P is the change of basis mat from the standard basis to our ON basis of evecs $v_1, \ldots, v_n$; A is of course the mat of the endomorphism wrt the standard basis.

L25

Let ψ a symmetric (herm) bilinear form on a real (cplx) vec sp V; let $A = [\psi]_B$ for some ON basis B of V, then $s(\psi) =$ no. $+ve$ evals of $A-$ no. $-ve$ evals of A: A is symmetric (herm) so by the above $\exists$ an ON B s.t. $P^{-1}AP = D = P^TAP$ $(\overline{P^T}AP)$ w/ D diagonal. But then the evals of D are those of $P^{-1}AP$ so we are done.

T26

Simultaneous diagonalization of quadratic forms: let ψ, ϕ symmetric (Herm) bilinear forms on a real (cplx) vecsp V; assume one of them, wlog ψ, is +ve definite (see Exs4Q10 for why this is actually necessary necessary), then $\exists$ basis B of V st $[\psi]_B, [\phi]_B$ diagonal: fix any basis and have mats A, C representing ψ, ϕ. Diagonalise ψ: $\exists$ non-singular mat P s.t. $P^TAP = I$ since ψ is +ve definite, now P^TCP is symmetric so m\exists an orthog mat R w/ $R^TP^TCPR = D$ diagonal, then $(PR)^T A(PR) = R^TIR = I$ and $(PR)^T C(PR) = D$ as above, and PR is nonsingular since P, R are, so we are done.

Rk

The diagonal entries of D are precisely the roots of the poly $\det(C - tA)$ since they are the roots of $\det(D - tI) = \det\left((PR)^T CPR - t(PR)^T APR\right) = \det(PR)^T \det(C - tA)\det(PR) = (\det PR)^2 \det(C - tA)$; since PR is non-singular the roots of this are precisely those of $\det(C - tA)$ as required.

Exercise: a symmetric mat is +ve definite iff the n principal minors (dets of submats in the top left corner of size $1, 2, \dots$) are +ve.

Final Rk

In IA A&G we looked at conics. For $n = 2$: $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0$. This is the locus of $\vec{x}^T A\vec{x} + B\vec{x} + C = 0$ where A is the symmetric mat $\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. For general n these [kinds of forms?] are called quadrics. Assume the conics are non-degenerate (not just points and straight lines) and we have the following cases:

- $s(A) = 2$, an ellipse. If we diagonalise using an orthog transformation $a_{11}x_1^2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0$ for new constants, then translate $x_i \to x_i - \frac{b_i}{2a_i}$ so $b_1 = b_2 = 0$ i.e. $a_{11}x_1^2 + a_{22}x_2^2 = c$; we can also squash by a non-orthogonal transform matrix P to $x_1^2 + x_2^2 = 1$, the unit circle.
- $r(A) = 2, s(A) = 0$; we similarly obtain a hyperbola $a_{11}x_1^2 - a_{22}x_2^2 = c$. On a 1D subsp the restricted form is +ve for lines between the two asymtopes in the same sections where the hyperbola is, and -ve for lines in the other two sections.
- $r(A) = s(A) = 1$: $a_{11}x_1^2 + b_1x_1 + b_2x_2 + c = 0$; translating we cannot eliminate b_2x_2 but have $a_{11}x_1^2 + b_2x_2 + c = 0$, then let $x_2 \to x_2 - \frac{c}{b_2}$ and $a_{11}x_1^2 + b_2x_2 = 0$, a parabola.

For $n > 2$ we get similar sets of cases, e.g. for $n = 3, r(A) = 3$:

- $s(A) = 3$: squashed form $x^2 + y^2 + z^2 = 1$, an ellipsoid
- $s(A) = 1$: squashed form $x^2 + y^2 - z^2 = 1$, a hyperboloid of one sheet
- $s(A) = -1$: squashed form $x^2 - y^2 - z^2 = 1$, a hyperboloid of two sheets

This concludes this course. For further reading and course is the next term, the lecturer recommends M Artin's "Algebra".