Miscellaneous

0.1 Basis

- The full set $z_i \otimes \zeta^j$ over all (i,j) is a basis for $\mathcal{L}(V,V)$.
- The full set $\zeta^i \otimes \zeta^j$ over all (i,j) is a basis for $\mathcal{L}(\hat{V},\hat{V})$. Note that this tensor product only takes in one vector.

0.2 Trace

- Invariant under similarity transformation (shown in Tut 6 Q6)
- Is sum of eigenvalues (shown in Tut 8 Q3)

0.3 How a matrix describes change Let T be a 2D LT.
$$\mathcal{M}(T) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
This means that $T(\hat{i}) = \begin{pmatrix} a \\ c \end{pmatrix}$ and $T(\hat{j}) = \begin{pmatrix} b \\ d \end{pmatrix}$, i.e. the columns tell us how the basic unit vectors change under the transformation.

0.4 Transpose of LT Let $\alpha \in \hat{V}$ and let $v \in V$. For each LT $T: V \to V$, define $\hat{T}: \hat{V} \to \hat{V}$ as follows:
$$\hat{T}(\alpha)(v) = \alpha(Tv)$$
Then $\hat{T}$ is the transpose of T. Since $\hat{T}$ is linear, so $\hat{T} \in \mathcal{L}(\hat{V},\hat{V})$.

0.5 Notation

- $I^i_j = 1$ if $i = j$, else 0 (known as Kronecker delta)
- Summation over pairs of subscript/superscript
- For a matrix M, the entry in row i and column j is M^i_j.

Isomorphisms

0.6 Definitions Let F be a mapping $F: S \to T$.

- Surjection: $\forall t \in T,\ \exists s \in S \quad F(s) = t$
- Injection: $\forall s_1, s_2 \in S \quad F(s_1) = F(s_2) \Rightarrow s_1 = s_2$
- Bijection: Surjection and Injection

0.7 Homomorphism Let $\phi: U \to V$ be a mapping. It is a homomorphism if
$$\phi(u+v) = \phi(u) + \phi(v)$$
$$\phi(au) = a\phi(u)$$
If this homomorphism is <u>also</u> a bijection, then this is an isomorphism.

0.8 Note LI $\Rightarrow$ injective $\Rightarrow$ surjective $\Rightarrow$ bijective $\Rightarrow$ basis

Eigenstuff

4.1 Linear operator A LT from V to itself

Complex operator A linear operator on a vector space over the complex numbers

4.2 Invariant subspace A subspace $U \subseteq V$ is invariant wrt $T \in \mathcal{L}(V,V)$ if $Tu \in U, \forall u \in U$.

4.7 Eigenray A 1-d invariant subspace. Its elements are eigenvectors.

4.10 Eigenvector/value Let T be a linear operator on V. A vector $v \in V$ is an eigenvector of T if $Tv = \lambda v$ for a particular $\lambda \in \mathcal{F}$. λ is the eigenvalue of this eigenvector.

4.19 Eigenrays are LI Consider N eigenrays with different eigenvalues. The set that consists of 1 vector from each eigenray is LI. (proof by contradiction)

4.20 Cap on number of eigenvalues The number of distinct eigenvalues for an operator V cannot be larger than $\dim(V)$. (k eigenvalues $\Rightarrow$ k eigenrays $\Rightarrow$ k LI vectors, then show injective linear map from $\mathcal{F}^n$ to V)

FTOA

Complex coefficients Any polynomial with complex coefficients can be completely factorized into the form
$$a(x-a_1)(x-a_2)\cdots$$
where $a, a_i \in \mathbb{C}$. The factorization is unique (excluding reordering)

Real coefficients Any polynomial with real coefficients can be expressed in the form
$$a(x-a_1)(x-a_2)\cdots(x^2+b_1x+c_1)(x^2+b_2x+c_2)\cdots$$
where $a, a_i, b_i, c_i \in \mathbb{R}$, and where each $b_i^2 < 4c_i$. The factorization is unique (excluding reordering).

4.21 Existence of an eigenvalue Every <u>complex operator</u> has at least one eigenvalue.

Proof

1. Let $\dim(V) = n$. Let $v \neq 0 \in V$.
2. The list of vectors $v, Tv, T^2v, \cdots, T^nv$ contains $n+1$ vectors, so it is not linearly independent.
$$\begin{aligned} c_0v + c_1Tv + \cdots + c_nT^nv &= 0 \\ (c_0I + c_1T + \cdots + c_nT^n)v &= 0 \quad \text{by linearity} \\ a(T-a_1I)(T-a_2I)\cdots v &= 0 \quad \text{by FTOA} \end{aligned}$$
3. If every $(T - a_jI)$ injective, then $v = 0$, contradiction.
4. Hence, some $(T - a_jI)$ is not injective, i.e. $(T - a_jI)w = 0$ has infinite solutions.
5. Hence a_j is an eigenvalue.

UT matrices

4.25 UT A square matrix M^i_j is upper triangular if $M^i_j = 0$ when $i > j$.

4.27 Similar Matrices A and B are similar if $A = P^{-1}BP$, where A, B are square and of the same size.

4.27A Similar matrices have the same eigenvalues Let M be a square matrix, and let P be a change-of-basis matrix. Let v be an eigenvector of M, corresponding to eigenvalue λ.
$$Mv = \lambda v$$
$$P^{-1}Mv = \lambda P^{-1}v$$
$$P^{-1}M(PP^{-1})v = \lambda P^{-1}v$$
$$(P^{-1}MP)(P^{-1}v) = \lambda(P^{-1}v)$$
Hence they have the same eigenvalues, but the eigenvectors are now $P^{-1}v$.

4.26 Every <u>complex operator</u> has a UT matrix. Or, every <u>complex matrix is similar</u> to a UT matrix.

Idea Let N be a complex matrix.

1. Choose arbitrary basis such that the first basis vector is an eigenvector.
$$P^{-1}NP = \begin{pmatrix} \lambda_1 & \alpha_1 & \alpha_2 \\ 0 & M^2_2 & M^2_3 \\ 0 & M^3_2 & M^3_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \alpha \\ 0 & M \end{pmatrix}$$
where M is 2 by 2 and α is 1 by 2. Note that λ_1 is an eigenvalue of N.
2. Do the same for M:
$$Q^{-1}MQ = \begin{pmatrix} \lambda_2 & \beta \\ 0 & R \end{pmatrix}$$
where $\beta, R \in \mathbb{C}$. Note that λ_2 is an eigenvalue of M (also of N).
3. But M is inside N, can we claim that we can apply a similarity transformation?
$$\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 & \alpha \\ 0 & M \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & \alpha Q \\ 0 & Q^{-1}MQ \end{pmatrix}$$
so we can concretely apply the similarity transformation by using $\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$.
4. By 3, we can change the basis of M by applying a similarity transformation to the original matrix. By 4.27A, then λ_2 is indeed an eigenvalue of N, so the claim in point 2 is true.

4.30 The diagonal entries in a UT matrix are its eigenvalues.

Proof Let M be a UT matrix.

- Consider $M - \lambda I$. It is UT, and by definition of λ, $M - \lambda I$ must be singular.
- Hence, we know one of the diagonal entries are 0, so since we subtracted λ, then one of the diagonal entries of M is λ.

Diagonal matrices

4.31 Diagonal A square matrix M^i_j is diagonal if $M^i_j = 0$ when $i \neq j$.

4.30A Since a diagonal matrix is also UT, its diagonal entries are its eigenvalues.

4.32-4.34 Diagonalizable An operator is diagonalizable if

- there is a basis wrt which its matrix is diagonal.
- or, there is a basis consisting of only eigenvectors
- or, the operator has $\dim(V)$ distinct eigenvalues

Eigenspaces

4.35 Eigenspace Let T be an operator on V. Then the eigenspace of T corresponding to eigenvalue λ is
$$E(\lambda, T) = Ker(T - \lambda I)$$
4.38 Eigenspaces are mutually disjoint The sum of eigenspaces corresponding to distinct eigenvalues is a <u>direct sum</u>, i.e.the intersection of those eigenspaces is a set containing the zero vector.

4.39 T is diagonalizable $\iff$
$$V = E(\lambda_1, T) \oplus E(\lambda_2, T) \oplus \cdots E(\lambda_k, T)$$
for distinct eigenvalues λ_i. (each eigenspace has a basis of eigenvectors)

Jordan canonical form

Superdiagonal of a square matrix is the set of elements directly above the diagonal. In the example below, the superdiagonal is exactly the set of all the 1s.

4.41 Jordan block A Jordan block of size m is a order m square matrix of the form (e.g. $m = 4$):
$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$
where the diagonal entries are a specific λ, the entries along the superdiagonal are 1, and the other entries are 0.

The Jordan block has only one eigenvector.

4.42 Jordan basis A Jordan basis is one such that the matrix of T consists of Jordan blocks:
$$\begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix}$$
Note that the eigenvalues of this matrix are the diagonal entries in the blocks, since the blocks are UT and so is the matrix.

4.43 Every <u>complex operator</u> has a unique Jordan basis (excluding reordering). (real operators may not have, because it might have complex eigenvalues)

4.44 Jordan canonical form When a operator is represented by the matrix wrt a Jordan basis, we say that it is in its JCF.

Note Real operators also have a JCF, but not as nice.

Finding change-of-basis matrix for JCF (Tut 7 Q8) Consider the shearing matrix S. Because it is UT we know its eigenvalues, thus we can construct its JCF:
$$S = \begin{pmatrix} 1 & \tan\theta \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
From 0.1, we know that the full set $z_i \otimes \zeta^j$ is a basis for $\mathcal{L}(V,V)$, so we can write both matrices in terms of their own bases:
$$S = e_1 \otimes \epsilon^1 + (\tan\theta) e_1 \otimes \epsilon^2 + e_2 \otimes \epsilon^2$$
$$J = \bar{e}_1 \otimes \bar{\epsilon}^1 + \bar{e}_1 \otimes \bar{\epsilon}^2 + \bar{e}_2 \otimes \bar{\epsilon}^2$$
Here, the bars represent the new basis under the similarity transformation that creates the JCF.

Notice that $\tan\theta e_1$ becomes $\bar{e}_1$. Thus, we can define $\bar{e_1} = \tan(\theta)e_1$. Then to ensure that duality is preserved after the transformation, we must have $\bar{\epsilon}^1 = \cot(\theta)\epsilon^1$. Notice we can leave e_2 and ϵ^2 as they are. So we have
$$\bar{e}_1 = (\tan\theta)e_1$$
$$\bar{e}_2 = e_2$$
and we can represent this in a matrix by 0.3:
$$\begin{pmatrix} \tan\theta & 0 \\ 0 & 1 \end{pmatrix}$$
and this is indeed the change-of-basis matrix for the JCF.

Cayley-Hamilton theorem

4.45 Multiplicty of an eigenvalue λ is

- the sum of the sizes of the Jordan blocks corresponding to the eigenvalue
- or, the number of times it appears along the diagonal

Diagonalizability of a matrix From Tut 7 Q7,

- The dimension of an eigenspace is the number of Jordan blocks of an eigenvalue
- A matrix is diagonalizable $\iff$ every eigenvalue has an eigenspace has max dimensionality, namely the multiplicity of that eigenvalue.

4.46 Characteristic polynomial Let m_i be the multiplicity of eigenvalue λ_i of a linear operator T. Then the characteristic polynomial of T is

$$\chi_T(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots$$

Clearly the eigenvalues satisfy $\chi_T(\lambda) = 0$.

Nilpotent matrix A square matrix M is nilpotent if $M^k = 0$ for some positive integer k, known as the index of M.

4.47 Cayley-Hamilton states that $\chi_T(T) = 0$.

Proof

1. Consider a Jordan block $B = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$, of some larger matrix M.
2. The multiplicity of λ_1 is 2, so there is a $(x-\lambda_2)^{m_2}$ term in the characteristic polynomial.
3. Consider $B - \lambda I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Note that it is nilpotent with index 2.
4. Thus $(B-\lambda I)^2 = 0$. This applies to all Jordan blocks. Take the product and we have the Cayley-Hamilton theorem.

4.50 Finding an inverse If a matrix M is invertible, then we can keep multiplying M^{-1} to its characteristic equation until there is a M^{-1} term, and make M^{-1} the subject of the formula. In particular, M^{-1} is a linear combination of M^i.

4.51 Geometric meaning of JCF Note that in the case with real nonzero eigenvalues:

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} 1 & 1/\lambda & 0 \\ 0 & 1 & 1/\lambda \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} 1 & \tan\phi & 0 \\ 0 & 1 & \tan\phi \\ 0 & 0 & 1 \end{pmatrix}$$

So every Jordan block describes a combination of stretching and shearing. As for eigenvalue 0, it represents a projection (into a smaller space). As for complex eigenvalues, it can include rotations.

Since every LT can be described by a collection of Jordan blocks, then every LT is a combination of stretching, projecting, shearing, and rotating.

Multilinear algebra

6.1 Multilinear form A multilinear form of degree m on an n-dimensional vector space V is a mapping

$$\underbrace{V \times V \times V \times \cdots}_{m \text{ times}} \to \mathcal{F}$$

which is linear in every slot.

6.1A Vector space The set of multilinear forms of degree m on V is a vector space in the usual way.

2-forms

6.2 2-form A bilinear form ψ on a real vector space V is called a 2-form if it is antisymmetric:

$$\psi(u,v) = -\psi(v,u)$$

so the space of 2-forms is a subset of $B(V)$. By 6.1A, the set of 2-forms on V is a vector space.

6.3 Wedge product Let $\alpha, \beta \in \hat{V}$. The wedge product of α and β is defined as:

$$\alpha \wedge \beta = \frac{1}{2}(\alpha\otimes\beta - \beta\otimes\alpha)$$

and it is a 2-form.

- It is indeed antisymmetric. For $u, v \in V$:
$$\begin{aligned}(\alpha\wedge\beta)(u,v) &= \frac{1}{2}\Big(\alpha(u)\beta(v) - \beta(u)\alpha(v)\Big) \\ &= -\frac{1}{2}\Big(\alpha(v)\beta(u) - \beta(v)\alpha(u)\Big) \\ &= -(\beta\wedge\alpha)(v,u)\end{aligned}$$
- Hence $\alpha\wedge\alpha = -\alpha\wedge\alpha$, i.e. $\alpha\wedge\alpha = 0$.

6.4 Basis Let ζ be the dual basis to the basis z for V.

- The full set $\zeta^i \wedge \zeta^j$ is a basis for the space of 2-forms on V.
- Any 2-form ψ can be written as $\psi = \psi_{ij}\zeta^i\wedge\zeta^j$.

2-forms on 3-dimensional vector space For a real 3-dimensional vector space with dual basis $\zeta^1, \zeta^2, \zeta^3$, the basis vectors of 2-forms are

$$\zeta^1\wedge\zeta^2, \quad \zeta^2\wedge\zeta^3, \quad \zeta^3\wedge\zeta^1$$

There are only 3 because:

- $\zeta^i\wedge\zeta^i = 0$ so they disappear
- $\zeta^i\wedge\zeta^j = -\zeta^j\wedge\zeta^i$ so they are linearly dependent

so the dimension of the space of 2-forms on a three dimensional space is 3. In particular, it is C^3_2.

3-forms

Definition A trilinear form ψ on a real vector space V is called a 3-form if it is antisymmetric in all 3 slots, e.g.

$$\psi(u,v,w) = -\psi(v,u,w) = -\psi(u,w,v)$$

Swapping any pair of vectors will negate the result.

Wedge product Let $\alpha, \beta, \gamma \in \hat{V}$. The wedge product of α, β and γ is defined as:

$$\begin{aligned}\alpha\wedge\beta\wedge\gamma = &\frac{1}{6}(\alpha\otimes\beta\otimes\gamma - \alpha\otimes\gamma\otimes\beta \\ &+ \beta\otimes\gamma\otimes\alpha - \beta\otimes\alpha\otimes\gamma \\ &+ \gamma\otimes\alpha\otimes\beta - \gamma\otimes\beta\otimes\alpha)\end{aligned}$$

If there are an odd number of swaps between vectors, then we put a minus sign. Again, this is a 3-form.

Basis Let ζ be the dual basis to the basis z for V. Then

- The full set $\zeta^i\wedge\zeta^j\wedge\zeta^k$ is a basis for the space of 3-forms on V.
- Any 3-form ψ can be written as $\psi = \psi_{ijk}\zeta^i\wedge\zeta^j\wedge\zeta^k$.
- The dimensionality of the space of 3-forms on an n-dimensional real vector space is C^n_3.

n-forms

For this section, let V be a n-dimensional vector space.

6.5 Dimensionality The space of m-forms on V is a vector space of dimension C^n_m.

6.6 The space of n-forms on V is a vector space of dimension C^n_n, i.e. it is 1-dimensional.

6.6A Transpose mapping Let T be a linear operator on V. Then define its transpose $\hat{T}$ on n-forms on V by

$$\hat{T}(\psi)(u,v,w,\cdots) = \psi(Tu,Tv,Tw,\cdots)$$

Then $\hat{T}(\psi)$ is also a n-form.

6.6B Repeated inputs If a n-form has repeated inputs, then it evaluates to 0. (Because of the antisymmetry property, and we can perform a "swap" among the repeated inputs)

$$\psi(u,u,\cdots) = -\psi(u,u,\cdots) \iff 2\psi(u,u,\cdots) = 0$$

$\Delta(T)$

6.7 Definition Let Ω be a non-zero n-form on V. Let T be a linear operator on V. By 6.6A, $\hat{T}\Omega$ is an n-form. By 6.6, the space of n-forms on V is 1-dimensional. Hence,

$$\hat{T}\Omega = \Delta(T)\Omega$$

where $\Delta(T)$ is a scalar.

6.8 Product $\Delta(TS) = \Delta(T)\Delta(S)$

Properties of $\Delta(T)$ for matrices

1. $\Delta(I) = 1$ (obvious when considering 6.6A definition)
2. Δ is invariant under a similarity transformation, since
$$\begin{aligned}\Delta(P^{-1}MP) &= \Delta(P^{-1})\Delta(M)\Delta(P) \\ &= \Delta(P^{-1}P)\Delta(M) = \Delta(M)\end{aligned}$$
3. Let M be a order n square matrix. Then
$$\Delta(M) = \Delta(PJP^{-1}) = \Delta(J)$$
where J is the Jordan canonical form.
4. For a Jordan block of size s, $\Delta(J) = \lambda^s$, since
$$\Omega(Je_i) = \Omega(\lambda e_i) = \lambda^s\Omega(e_i)$$
(LHS is $\hat{\Omega}$ because it is the RHS of 6.6A)
5. (6.9) Let M have JCF J. Then
$$\Delta(M) = \Delta(J) = \lambda_1^{m_1}\lambda_2^{m_2}\cdots$$

Properties of $\Delta(T)$ for LTs

1. (6.10) $\Delta(T) = \Delta(\mathcal{M}(T))$
2. (6.11) $\Delta(T) = \lambda_1^{m_1}\lambda_2^{m_2}\cdots$, by 6.9
3. (6.12) $\Delta(T)\neq 0 \iff \lambda\neq 0 \iff T$ is bijective.
4. (6.13) $\Delta(T-\lambda I) = 0$ (because $T-\lambda I$ must not be bijective)
5. $\Delta(T)$ is the determinant of T. It is the eigenvalue of $\hat{T}$ on a 1-dimensional vector space.

Proof for 6.10 Let $z : \mathcal{F}^n \to V$ be a basis on V. Let Ω be a n-form on V. Let $a, b, \cdots$ be vectors in $\mathcal{F}^n$. Define a non-zero n-form Ω^* on $\mathcal{F}^n$:

$$\Omega^*(a,b,\cdots) = \Omega(za,zb,\cdots)$$

Let $T : V \to V$ be a LT.

$$\begin{aligned}&\Delta(T)\Omega(za,zb,\cdots) \\ &= \hat{T}\Omega(za,zb,\cdots) && \text{by 6.7} \\ &= \Omega(Tza,Tzb,\cdots) && \text{by 6.6A} \\ &= \Omega^*(z^{-1}Tza,z^{-1}Tzb,\cdots) && \text{by } \Omega^* \text{ defn} \\ &= \Omega^*(\mathcal{M}(T)a,\mathcal{M}(T)b,\cdots) && \text{by Fact 3.65} \\ &= \Delta(\mathcal{M}(T))\Omega^*(a,b,\cdots) && \text{by 6.6A and 6.7} \\ &= \Delta(\mathcal{M}(T))\Omega(za,zb,\cdots) && \text{by } \Omega^* \text{ defn}\end{aligned}$$

Since $a, b, \cdots$ are arbitrary and $\Omega\neq 0$, then

$$\Delta(T) = \Delta(\mathcal{M}(T))$$

Every square matrix is similar to its transpose Let $A = PDP^{-1}$.

$$A^{\mathrm{T}} = (PDP^{-1})^{\mathrm{T}} = (P^{-1})^{\mathrm{T}}D^{\mathrm{T}}P^{\mathrm{T}} = (P^{\mathrm{T}})^{-1}DP^{\mathrm{T}}$$

Other properties

- Because every square matrix is similar to its transpose, by property 1 for matrices,
$$\Delta(M) = \Delta(P^{-1}MP) = \Delta(M^{\mathrm{T}})$$
- The determinant switches sign when swapping two rows/columns in the original matrix, because n-forms are antisymmetric in all pairs of slots.
- We can defined a generalized parallelopiped, formed using the eigenvectors of LT T. Then, under T, its volume changes by a factor given by the product of the stretching factors in all directions, which is the product of the stretching factors for all eigenvectors, which is precisely $\lambda_1^{m_1}\lambda_2^{m_2}\cdots$.

Triple product on $\mathbb{R}^3$ For $u, v, w \in V$, the vector triple product is the map

$$(u,v,w) \to u\cdot v\times w$$

It is trilinear and antisymmetric in all slots, so it is a 3-form on $\mathbb{R}^3$. Since 3 is both in the 3-form and the dimensionality of $\mathbb{R}^3$, then by 6.7,

$$(u,v,w) \to (Tu)\cdot(Tv)\times(Tw) = \Delta(T)(u\cdot v\times w)$$

is another 3-form, which is proportional to the triple product. This was how we defined the determinant in $\mathbb{R}^3$ in Chapter 1.

Note: the triple product on $\mathbb{R}^3$ is defined using the lengths of vectors and the angles between them, so it depends on the choice of *dot* as the inner product. It is called the volume form associated with *dot*.

Tensors

6.14 Let V be a vector space. Let $R : V\times V \to \mathcal{L}(V,V)$. R is called a tensor.

Lengths and angles

5.1 Dot product on $\mathbb{R}^n$ Let $u = u^i e_i$ and $v = v^j e_j$, where $u, v \in \mathbb{R}^n$. Then $u \cdot v = \sum_{k=1}^{n} a^k b^k$, i.e. we take the sum of the products elementwise.

- Length: $|u| = \sqrt{u \cdot u}$
- Components: $e_k \cdot v = v^k$
- Cosine of angle: $\cos(\theta) = \dfrac{u \cdot v}{|u|\,|v|}$

Bilinear forms

5.2 Bilinear form A bilinear form on a vector space V is a mapping $b : V \times V \to \mathbb{R}$ which is linear in both slots:

$$b(cu + dv, w) = cb(u, w) + db(v, w)$$
$$b(w, cu + dv) = cb(w, u) + db(w, v)$$

for $u, v, w \in V$ and $c, d \in \mathbb{R}$.

Note: not necessary that $b(u, v) = b(v, u)$.

5.4 The set of bilinear forms on V, $B(V)$, is a vector space in the usual way.

5.5 Tensor product Let $\alpha, \beta \in \hat{V}$, and let $u, v \in V$. Then the tensor product of α with β is the element of $B(V)$ such that

$$\alpha \otimes \beta(u, v) = \alpha(u)\beta(v)$$

5.8, 5.9 Basis for $B(V)$ Let z_i be a basis for V and let ζ^j be the dual basis.

- For each (i, j), $\zeta^i \otimes \zeta^j \in B(V)$.
- The full set $\zeta^i \otimes \zeta^j$ over all (i, j) is a basis for $B(V)$.
- The dot product on $\mathbb{R}^n$ is an element of $B(V)$.

Proof for span

1. Let $b \in B(V)$ and define $b_{ij} = b(z_i, z_j)$.
2. Define $h = b_{ij}\zeta^i \otimes \zeta^j$, and let it act on (z_k, z_l):
$$h(z_k, z_l) = b_{ij}\zeta^i \otimes \zeta^j(z_k, z_l) = b_{ij}\zeta^i(z_k)\zeta^j(z_l)$$
$$= b_{ij} I^i_k I^j_l = b_{kl} = b(z_k, z_l)$$
By linearity, then $h(u, v) = b(u, v)$ for all $u, v \in V$.
3. Thus, $h = b$, that is $b = b_{ij}\zeta^i \otimes \zeta^j$, so the full set of $\zeta^i \otimes \zeta^j$ span $B(V)$. We call the numbers b_{ij} the components of b relative to basis z_i.

5.11 Extracting components
Since $b_{ij} = b(z_i, z_j)$, we just let b act on (z_i, z_j) to get the components of b.

5.13 Since dot is a bilinear form, we can express it as a linear combination of $\zeta^i \otimes \zeta^j$, by first finding its components (relative to the chosen basis e_i):

$$b_{ij} = dot(e_i, e_j) = e_i \cdot e_j = I_{ij}$$

and by observation, $dot = I_{ij}\epsilon^i \otimes \epsilon^j$. Note that we must balance the subscripts and superscripts.

5.14 Matrix of bilinear form The matrix of a bilinear form $b \in B(V)$, where $\dim(V) = 3$, rel. to basis z_i is:

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

5.16 Change of basis From Tut 6 Q7, if we change to a new vector basis $y_i = P^j_i z_j$, then the corresponding dual bases are related by $\eta^j = (P^{-1})^j_i \zeta^i$, so $\zeta^j = P^j_i \eta^i$. Thus,

$$b = b_{ij}\zeta^i \otimes \zeta^j = b_{ij} P^i_k \eta^k \otimes P^j_h \eta^h = (P^i_k b_{ij} P^j_h)\eta^k \otimes \eta^h$$

Refer to 5.77 to find out more

Inner products

5.18 Inner product space is a pair (V, g), with a vector space V, and a particular inner product g, which is a bilinear form with the following properties:

1. Positivity: $g(u, u) \geq 0 \quad \forall u \in V$
2. Definiteness: $g(u, u) = 0 \iff u = 0$
3. Symmetry: $g(u, v) = g(v, u) \quad \forall u, v \in V$

5.22 The set of inner products on V is a subset of $B(V)$. But it is not a subspace, because it does not contain the zero bilinear form, as it does not satisfy definiteness.

5.24, 5.25 Consider a bilinear form b. Consider its matrix b_{ij} relative to basis z_i. If b_{ij} is diagonal, and its diagonal entries $b_{ii} > 0$, then

$$b = b_{ij}\zeta^i \otimes \zeta^j$$

is an inner product, where ζ^i is the dual basis to z_i.

5.26 Properties of inner product matrix Let $u^i \in \mathbb{R}^n$. The matrix g_{ij} of the bilinear form g rel. to basis z_i is $(n \times n)$ matrix with the following analogous properties:

1. Positivity: $g_{ij}u^i u^j \geq 0$
2. Definiteness: $g_{ij}u^i u^j = 0 \iff u^i = 0 \quad \forall i$
3. Symmetry: $g_{ij} = g_{ji} \quad \forall i, j$

Definiteness implies non-singular

- By definition, a singular matrix g has $g_{ij}u^j = 0$ where $u \neq 0$.
- Multiplying by u^i and summing, we have $g_{ij}u^i u^j = 0$.
- By definiteness, the above equation implies $u^i = 0$ for all i, i.e. $u = 0$, contradicting our assumption.
- Hence g_{ij} is non-singular.

Lengths and angles

For this section, let (V, g) be an inner product space.

5.29 Length The length (or norm) of a vector $v \in V$ is

$$|v| = \sqrt{g(v, v)}$$

5.30 Orthogonal Vectors $u, v \in V$ are orthogonal if $g(u, v) = 0$.

5.31 If u, v are orthogonal, then $|u + v|^2 = |u|^2 + |v|^2$.

5.33 Orthogonal decomposition Let $u, v \in V$ where $v \neq 0$. Then $u = u_{\parallel v} + u_{\perp v}$, where

$$u_{\parallel v} = \frac{g(u, v)}{|v|^2}v \quad \text{and} \quad u_{\perp v} = u - \frac{g(u, v)}{|v|^2}v$$

5.34 Cauchy-Schwarz inequality If $u, v \in V$, then

$$|g(u, v)| \leq |u|\,|v|$$

Equality occurs if and only if $u = kv$ for scalar k.

Proof

1. If $v = 0$, $g(u, 0) = 0$, and RHS $= 0$, so the statement is true.
2. Consider $v \neq 0$, applying 5.31 to the orthogonal decomposition:
$$|u|^2 = |u_{\parallel v}|^2 + |u_{\perp v}|^2 \qquad (1)$$
3. Since (1) is a sum of squares, we have $|u|^2 \geq |u_{\parallel v}|^2$.
4. Then, compute $u^2_{\parallel v}$:
$$|u_{\parallel v}|^2 = g\left(\frac{g(u, v)}{|v|^2}v, \frac{g(u, v)}{|v|^2}v\right) = \left(\frac{g(u, v)}{|v|^2}\right)^2 g(v, v)$$
$$= \left(\frac{g(u, v)}{|v|^2}\right)^2 |v|^2 = \frac{g(u, v)^2}{|v|^2}$$
5. Susbtituting into the inequality $|u|^2 \geq |u_{\parallel v}|^2$, we get
$$|u|^2 \geq \frac{g(u, v)^2}{v^2}$$
Rearranging and taking the positive square root, we get the Cauchy-Schwarz inequality.
6. Referring to (1), we get equality if $|u_{\perp v}| = 0$, i.e. $u_{\perp v} = 0$. That is, u and v are parallel.

5.35 Triangle inequality For $u, v \in V$,

$$|u + v| \leq |u| + |v|$$

Equality occurs if and only if $u = kv$ for $k \geq 0$.

Proof Again for $u = 0$ it is obvious. For $u \neq 0$,

$$|u + v|^2 = |u|^2 + |v|^2 + 2g(u, v) \leq |u|^2 + |v|^2 + 2\,|g(u, v)|$$
$$\leq |u|^2 + |v|^2 + 2\,|u|\,|v|$$
$$= (|u| + |v|)^2$$

For equality, we must first get equality from the Cauchy-Schwarz inequality, i.e. $v = ku$ for some scalar k. Also, we must satisfy $|g(u, v)| = |u|\,|v|$. Substituting, we must have $k \geq 0$.

5.36 Angle Let $u, v \in V$. The angle between u and v, θ is

$$\cos(\theta) = \frac{g(u, v)}{|u|\,|v|}$$

where $0 \leq \theta \leq \pi$. Note that this equation makes sense because of Cauchy-Schwarz, which ensures that the RHS lies in $[-1, 1]$.

Orthonormal bases

For this section, let (V, g) be an inner product space.

5.37, 5.38 Orthonormal basis Let z_i be a basis on V. It is an orthonormal basis if for all i, j, we have

$$g(z_i, z_j) = I_{ij}$$

or equivalently, its matrix relative to z_i is I.

5.40 The canonical basis for $\mathbb{R}^n$ is an orthonormal basis for $(\mathbb{R}^n, \cdot)$.

5.41 We can alternatively define an orthonormal basis z as follows:

$$g(za, zb) = dot(a, b)$$

where $a, b \in \mathbb{R}^n$.

5.42 Suppose y_i is a list of orthornormal vectors. Then

$$\left|a^i y_i\right|^2 = (a^1)^2 + (a^2)^2 + (a^3)^2 + \cdots$$

5.43 Orthonormal vectors are LI Any list of orthonormal vectors is LI. (square root both sides of 5.42, then equate to 0. Sum of squares is 0 $\Rightarrow$ each term is 0)

5.44 Any list of $n = \dim(V)$ orthonormal vectors is a basis for V.

Gram-Schmidt orthogonalization

5.45 Unit vector A unit vector v^+ is one where $|v^+| = 1$. For any vector u, we can scale it to construct a unit vector:

$$u^+ = \frac{u}{|u|}$$

5.46 Parallel component of unit vector If v is a unit vector, then for any vector u,

$$u_{\parallel v} = \frac{g(u, v)}{|v|^2}v = g(u, v)v$$

5.46A GS is an algorithm that transforms a basis into an orthonormal basis. Example for $\dim(V) = 3$:

$$z_1^+ = \frac{z_1}{|z_1|}$$
$$z_2^+ = \frac{z_2 - g(z_2, z_1^+)z_1^+}{\left|z_2 - g(z_2, z_1^+)z_1^+\right|}$$
$$z_3^+ = \frac{z_3 - g(z_3, z_1^+)z_1^+ - g(z_3, z_2^+)z_2^+}{\left|z_3 - g(z_3, z_1^+)z_1^+ - g(z_3, z_2^+)z_2^+\right|}$$

- Removes the parallel component of each new orthonormal basis vector.
- Each z_i^+ is a linear combination of $z_1, z_2, \cdots, z_i$, so the change-of-basis matrix is UT.

5.47 Every inner product space has an orthonormal basis.

Riesz representation

For this section, let (V, g) be an inner product space.

5.48 Riesz representation Let $u \in V$. Define a dual vector Γ_u:

$$\Gamma_u(v) = g(v, u) \quad \forall v \in V$$

It maps V to $\mathbb{R}$ and is linear because g is bilinear. Hence $\Gamma_u \in \hat{V}$.

5.49 Define a mapping $\Gamma : V \to \hat{V}$ such that

$$\Gamma : u \to \Gamma_u$$

Γ is linear because g is bilinear. Γ is an isomorphism.

Proof (Isomoprhism)

1. Assume Γ is not injective. Then some $\Gamma(v)$ for $v \neq 0$ is the zero dual vector.
2. Then $\Gamma(v)(u) = 0 \quad \forall u \in V$.
3. In particular, $\Gamma(v)(v) = 0$ for some $v \neq 0$, contradicting the definiteness of g, so Γ is injective.
4. By FT of LA, the $Rank(\Gamma) = \dim(V)$. But we know $\dim(V) = \dim(\hat{V})$.
5. Thus $Rank(\Gamma) = \dim(\hat{V})$ so the range of Γ is all of $\hat{V}$, so Γ is surjective.
6. Since Γ is linear (since g is linear in both slots), then Γ is an isomorphism.

5.50 Riesz representation Any dual vector α can be expressed as $\Gamma(u) = \Gamma_u$ for some unique $u \in V$.

$$\alpha(v) = \Gamma(u)(v) = g(v, u) \quad \forall v \in V$$

5.51 Dual basis of orthonormal basis For an orthonormal basis z_i, the dual basis is $\Gamma(z_j)$.

Proof

- Let $v = a^i z_i \in V$.
- By 5.50, $\Gamma(z_j)(z_i) = g(z_i, z_j) = I_{ij}$
- $\Gamma(z_j)v = \Gamma(z_j)(a^i z_i) = a^i \Gamma(z_j)(z_i) = I_{ij} a^i = a^j$
- The jth component of v is obtained by letting $\Gamma(z_j)$ act on v, i.e. it is a dual vector.

Riesz representation in terms of components

- Let z_i be a basis, let ζ^j be the dual basis.
- Let $\alpha = p_i \zeta^i, u = a^j z_j, v = b^k z_k$.
- Then $\alpha(v) = g(v, u)$ is written as
$$p_i \zeta^i b^k \zeta_k = p_k b^k = g(b^k z_k, a^j z_j) = a^j b^k g_{kj}$$
- Since this is true for arbitrary v (and thus arbitrary b^k, we must have $p_k = g_{kj} a^j$.
- This means that u can be found from α, and it asserts that matrix g_{kj} is not singular.

The complex case

Dot product on $\mathbb{C}^n$ Let $u = u^i e_i$ and $v = v^j e_j$, where $u, v \in \mathbb{C}^n$. Then $u \cdot v = \sum_{k=1}^{n} a^k \bar{b}^k$.

Properties of complex conjugates Let $z = a + bi \in \mathbb{C}$. Let $w \in \mathbb{C}$.

- $\bar{z} = \overline{a + bi} = a - bi$
- $\overline{z \pm w} = \bar{z} \pm \bar{w}$
- $\overline{(zw)} = \bar{z}\,\bar{w}$
- $\overline{(\bar{z})} = z$
- $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$
- z is real $\iff \bar{z} = z$.

Conjugate dual vectors (Tut 10 Q5) Let V be a complex vector space. Let $\gamma : V \to \mathbb{C}$ be a mapping which is conjugate linear, that is if $a, b \in \mathbb{C}$ and $u, v \in V$, then γ satisfies

$$\gamma(au + bv) = \bar{a}\gamma(u) + \bar{b}\gamma(v)$$

Also, if α is a dual vector and $v \in V$, then if we define $\bar{\alpha}$ as follows, it is a conjugate dual vector:

$$\bar{\alpha}(v) = \overline{\alpha(v)}$$

5.53 Sesquilinear form A sesquilinear form on a complex vector space V is a mapping $s : V \times V \to \mathbb{C}$ which is linear in the first slot, but conjugate-linear in the second:

$$s(cu, v) = cs(u, v)$$
$$s(u, cv) = \bar{c}s(u, v)$$

for $u, v \in V$ and $c \in \mathbb{C}$.

5.54 Complex inner product space is a pair (V, g), with a vector space V, and a particular sesquilinear form g that satisifes

1. Positivity: $g(u, u) \geq 0 \quad \forall u \in V$
2. Definiteness: $g(u, u) = 0 \iff u = 0$
3. Symmetry: $g(u, v) = \overline{g(v, u)} \quad \forall u, v \in V$

Other properties

- Matrix is same as in 5.14
- $g(u, u) \in \mathbb{R}$ by the symmetry property
- Length is same as in 5.29 (since $g(u, u) \in \mathbb{R}$)
- Orthogonality, Cauchy-Schwarz is same as in 5.30, 5.34

Triangle inequality For $u, v \in V$,

$$|u + v|^2 \leq |u|^2 + |v|^2 + 2\,|u|\,|v|$$

Proof Again for $u = 0$ it is obvious. For $u \neq 0$:

$$\begin{aligned}
|u + v|^2 &= g(u + v, u + v) = |u|^2 + |v|^2 + g(u, v) + g(v, u) \\
&= |u|^2 + |v|^2 + g(u, v) + \overline{g(u, v)} \\
&= |u|^2 + |v|^2 + 2\mathrm{Re}(g(u, v)) \\
&\leq |u|^2 + |v|^2 + 2\,|g(u, v)| \quad \text{since } \mathrm{Re}(c) \leq |c| \\
&\leq |u|^2 + |v|^2 + 2\,|u|\,|v| \quad \text{by Cauchy-Schwarz}
\end{aligned}$$

5.55 Riesz representation Let $u \in V$. Define a dual vector Γ_u:

$$\Gamma_u(v) = g(v, u) \quad \forall v \in V$$

In the complex case, it is not an isomorphism. Instead, it is a conjugate isomorphism.

5.56 Conjugate isomorphism Let V, W be complex vector spaces. A conjugate isomorphism $F : V \to W$ is a bijective map that satisfies

$$F(au + bv) = \bar{a}F(u) + \bar{b}F(v) \quad \forall a, b \in C, u, v \in V$$

5.57 Schur's theorem Given a linear operator T on a complex inner product space, there is a orthonormal basis wrt which T has a UT matrix.

Proof

1. By 4.26, every complex matrix is similar to a UT matrix.
2. We can use GS to get to an orthonormal basis. By 5.46A, the change-of-basis matrix is UT.
3. By Tut 9 Q5, the product of two UT matrices is UT, hence proven.

Spectral theorem

5.58 Bijection between LTs and sesquilinear/bilinear forms

By the Riesz representation, we can think of vectors as dual vectors. We can also think of LTs as sesquilinear forms as follows. (The same explanation applies if we define τ to be bilinear)

Let T be an operator on the complex inner product space (V, g). Then define a sesquilinear form τ:

$$\tau(u, v) = g(u, Tv) \quad \forall u, v \in V$$

To get the components,

$$\tau_{ij} = \tau(z_i, z_j) = g(z_i, Tz_j) = g(z_i, T_j^k z_k) = g_{ik}\overline{T_j^k}$$

Since g_{ij} is invertible, we can express

$$\overline{T_j^k} = g^{ik}\tau_{ij}$$

where g^{ik} is the inverse of g_{ik}, so that each τ defines a unique operator on (V, g).

5.59 Hermitian/Symmetric form A sesquilinear/bilinear form is Hermitian (in the complex case) or symmetric (in the real case) if

$$\tau(u, v) = \overline{\tau(v, u)} \quad \forall u, v \in V$$

5.60 Hermitian/Symmetric LT A LT on an inner product space is Hermitian/symmetric if its corresponding sesquilinear/bilinear form is Hermitian/symmetric. In particular, T satisfies

$$g(u, Tv) = \tau(u, v) = \overline{\tau(v, u)} = \overline{g(v, Tu)} = g(Tu, v)$$

and the matrix of τ is Hermitian/symmetric:

$$\tau_{ij} = \tau(z_i, z_j) = \overline{\tau(z_j, z_i)} = \overline{\tau_{ji}}$$

in the complex case, and $\tau_{ij} = \tau_{ji}$ in the real case.

5.63 If we use an orthonormal basis, then the matrices of T and τ are complex conjugates of each other. In particular, if they are real, then they are equal.

$$\tau_{ij} = g_{ik}\overline{T_j^k} = \overline{T_j^i}$$

because $g_{ik} = 1 \iff i = k$.

5.64 The eigenvalues of a Hermitian/symmetric LT are real.

5.65 Proof By 4.21, let T have an eigenvector u with eigenvalue λ. On one hand we have

$$g(u, Tu) = g(u, \lambda u) = \bar{\lambda}g(u, u) \tag{2}$$

on the other hand

$$g(Tu, u) = \lambda g(u, u) \tag{3}$$

By the definition in 5.60, LHS of (2) and (3) are equal. Thus, we have

$$\bar{\lambda}g(u, u) = \lambda g(u, u)$$

but since $u \neq 0 \Rightarrow g(u, u) \neq 0$ (definiteness), then

$$\bar{\lambda} = \lambda$$

i.e. λ is real.

5.66, 5.67 Spectral theorem Any Hermitian LT T on a complex inner product space has a real, diagonal matrix relative to some orthonormal basis, i.e. it can be diagonalized and the resulting matrix is real.

Proof

1. By 5.57, T can be made UT wrt an orthonormal basis.
2. By 5.63, the associated sesquilinear form is also UT.
3. By 5.61, $\tau_{ij} = \overline{\tau_{ji}}$. The UT entries are the complex conjugates of the lower-triangular (LT) entries.
4. Since the LT entries are 0, the UT entries are 0.
5. Hence the matrix is diagonal.

5.68 The spectral theorem also applies to symmetric LTs on real vector spaces.

5.69 The eigenvectors of a Hermitian transformation form an orthonormal basis. (The basis used is orthonormal. Since it can diagonalize the matrix, it is a basis of eigenvectors.)

Tut 10 Q1 The eigenvectors of an Hermitian transformation corresponding to distinct eigenvalues are orthogonal. (consider $g(u, Tv) = g(Tu, v)$, letting u and v be eigenvectors with distinct eigenvalues)

5.70 Spectral decomposition Let S be a real symmetric bilinear form, let z_i be the orthonormal basis of eigenvectors mentioned in 5.69, let λ_i be the eigenvalues, and let ζ^i be the dual basis to z_i. Since the matrix relative to the basis is diagonal, then we have

$$S = \lambda_1 \zeta^1 \otimes \zeta^1 + \lambda_2 \zeta^2 \otimes \zeta^2 + \cdots$$

Decomposing matrices

5.71 Transpose matrix The transpose of a (not necessarily square) matrix $M = M_{ij}$ is $M^{\mathrm{T}} = M_{ji}$.

5.72 Orthogonal matrix A real matrix P is orthogonal if

1. $P^{\mathrm{T}}P = I$, or
2. Rows/columns are all unit vectors, and rows/columns are orthogonal to other rows/columns wrt *dot* on $\mathbb{R}^n$.

5.76 Any symmetric matrix S can be expressed in the form $O^{\mathrm{T}}DO$, where D is diagonal and real, and O is orthogonal. We say that all symmetric real matrices are orthogonally diagonalizable.

Unitary matrix A complex matrix is unitary if

1. $\overline{U^{\mathrm{T}}}U = I$, or
2. Rows/columns are all unit vectors, and rows/columns are orthogonal to other rows/columns wrt *dot* on $\mathbb{C}^n$.

5.77 Any Hermitian matrix H can be expressed in the form $\overline{U^{\mathrm{T}}}DU$, where D is diagonal and real, and U is a complex unitary matrix. We say that all Hermitian matrices are unitarily diagonalizable.

Example Let S be a real, symmetric matrix.

1. Consider the inner product space $(\mathbb{R}^n, dot)$, which has canonical dual basis ϵ^i. By 5.24, we can define a bilinear form on $\mathbb{R}^n$. By 5.60, if the bilinear form S is symmetric, then the LT S is also symmetric. By 5.58, we have a bijection between bilinear forms and LTs so the statement of 5.60 is an if and only if. So indeed the bilinear form S is symmetric.
$$S = S_{ij}\epsilon^i \otimes \epsilon^j$$
2. Next, since we have a real symmetric bilinear form, by 5.70, we know that $\mathbb{R}^n$ has an orthonormal basis z_i wrt inner product *dot*, wrt which the matrix of S is a diagonal real matrix:
$$S = D_{ii}\zeta^i \otimes \zeta^i = D_{ij}\zeta^i \otimes \zeta^j$$
   - RHS is still equal because D is a diagonal matrix.
   - z_i is not the canonical basis e_i, because if it were, then the original matrix would have been diagonal.
3. Let P_j^i be the change-of-basis matrix from z_i to the canonical basis e_i, so $e_i = P_i^j z_j$. By 5.16, we have
$$S = (P_i^h D_{hk} P_j^k)\epsilon^i \otimes \epsilon^j$$
so $S_{ij} = P_i^h D_{hk} P_j^k$.
4. In terms of matrices, by 5.71, we have $S = P^{\mathrm{T}}DP$ for some matrix P.
5. Then, since z_i and e_i are both orthonormal bases wrt *dot*, and since $e_i = P_i^j z_j$, we have
$$dot(e_i, e_j) = I_{ij} = dot(P_i^h z_h, P_j^k z_k) = P_i^h I_{hk} P_j^k$$
i.e. $P^{\mathrm{T}}P = I$, so P is orthogonal.