MA2101 Test 2 Cheatsheet

by Devansh Shah

Chapter 1

Determinants and Inverses

The area of a parallelogram described by the 2 vectors $\vec{u}$ and $\vec{v}$ is given by $|\vec{u} \times \vec{v}|$.
Similarly, the volume of a parallelopiped described by $\vec{u}, \vec{v}, \vec{w}$ is given by $|\vec{u} \times \vec{v} \cdot \vec{w}|$.
Determinant of a transformation (in 2D) is defined to be the ratio of the final area of the basic box to the initial area of the basic box.
A **singular** linear transformation:

- maps two different vectors to one vector
- destroys all of the vectors in at least one direction
- loses all information associated with these directions
- satisfies $\det T = 0$

We can find the inverse by finding the transpose of the cofactor matrix, and dividing by the determinant of the original matrix.

Change of Basis

Let $\mathbf{P}$ be the transformation that takes $\hat{i} \to \vec{u}$ and $\hat{j} \to \vec{v}$ where $\vec{u}, \vec{v}$ are not parallel (assume 2-D). That is, $\vec{u} = P\hat{i}$ and $\vec{v} = P\hat{j}$.
Then, the components of ANY vector relative to $\vec{u}, \vec{v}$ are obtained by multiplying $\mathbf{P}^{-1}$ into the components relative to $\hat{i}, \hat{j}$.
And the matrix of any transformation $\mathbf{T}$ relative to $\vec{u}, \vec{v}$ is obtained by multiplying $\mathbf{P}^{-1}$ on the left and $\mathbf{P}$ on the right into the matrix of $\mathbf{T}$ relative to $\hat{i}, \hat{j}$.
The matrix of a transformation relative to its own eigenvectors (assuming that these form a basis, which is not always the case) is diagonal.
If the matrix $\mathbf{M}$ can be diagonalized, then $\mathbf{M} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is the diagonal matrix consisting of the eigenvalues of $\mathbf{M}$, and $\mathbf{P}$ is the matrix that transforms the original (standard) basis to the eigenvectors
Not every matrix can be diagonalized.
Trace of $\mathbf{M}$ denoted by TrM is the sum of the diagonal entries.
In general, $TrMN = TrNM$ but $TrMN \neq TrM\ TrN$.
Hence, $Tr(P^{-1}DP) = Tr(DPP^{-1}) = TrD$, so the trace of a matrix is always the same, no matter which basis you use.
The trace is equal to the sum of the eigenvalues. The determinant is equal to the product of the eigenvalues.
The trace and determinant of a linear transformation are well-defined, they don't depend on the basis chosen to represent the transformation in matrix form.

Chapter 2

Subspace

A subset U of a vector space V is called a **subspace** if U is a vector space with the same scalar multiplication and addition as in V.
Suppose U_1 and U_2 are sub-SETS of a vector space V. Then, the **sum** of these subsets, denoted by $U_1 + U_2$ is defined as the subset of V consisting of all vectors of the form $u_1 + u_2$ where $u_1 \in U_1, u_2 \in U_2$.
The sum of two *subspaces* is again a subspace.
Let U_1 and U_2 be *subspaces* of a vector space V. Suppose that the intersection $U_1 \cap U_2 = \{0\}$. Then, $U_1 + U_2$ is called the **direct sum** of U_1 and U_2, denoted by $U_1 \oplus U_2$.
Suppose that you've expressed a vector $v \in V$ as $u_1 + u_2 \in U_1 \oplus U_2$, then that expression is unique.

Spanning, Linear Independence, Basis

A basis for V is a vector space isomorphism (mapping) from $F^n \to V$, where n is the dimensionality of the vector space.
Suppose that $z : F^n \to V$ is a basis (in the new sense, i.e., it is a vector space isomorphism). Then, the list of vectors $z_i \equiv z(e_i)$ is a basis for V (in the old-fashioned sense, i.e., it is a list of vectors).
Suppose U_1 and U_2 are subspaces of a finite-dimensional vector space V, with the intersection consisting only of the zero vector. Then, $\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2)$.
Suppose U is a subspace of a finite-dimensional vector space V. Then, V has another subspace W such that $V = U \oplus W$.

Chapter 3

A mapping T is **linear** if:

- $T(u + v) = T(u) + T(v)\ \forall u, v \in V$
- $T(au) = aT(u) \forall a \in F,\ \forall v \in V$

The set of all linear mappings from V to W is denoted $L(V, W)$.
A basis of V is just an element of $L(F^n, V)$.
If z and y are bases (in the abstract sense, as mappings), then we can change the basis by multiplying z with $Q = y \circ z^{-1}$ so that:

$$Qz_i = Qz(e_i) = y \circ z^{-1} z(e_i) = y(e_i) = y_i$$

The set $\hat{V} = L(V, F)$ is called the **dual space** to the space V. An element of $\hat{V}$ is called a **dual vector**.
$L(V, W)$ is itself a vector space where:

- $(aT)(u) \equiv a(Tu)\ \forall v \in V$
- $(T_1 + T_2)(u) = T_1(u) + T_2(u)\ \forall u \in V$

Kernel and Range

If $T \in L(V, W)$, then Ker(T) (called **Kernel**) is the set of all elements in V that are mapped to the zero vector in W. It is a subspace of V.
The dimension of Ker(T) is called the **Nullity** of T, Null(T).
T is injective if and only if Ker(T) consists of the zero vector only.
The **range** of T is the set of vectors in W which can be expressed as Tv for some $v \in V$. Range(T) is a subspace of W.
The dimension of Range(T) is called the **rank** of T.
For any transformation T from V to W,

$$dim(V) = Null(T) + Rank(T)$$

For *linear* maps from V to W where $dim(V) = dim(W)$: bijectivity $\iff$ injectivity $\iff$ surjectivity.

Duality

If $v = a^i z_i$, then for each i, we have: $\zeta^i(v) = a^i$. So, ζ^i extracts the ith component of any vector, relative to the basis z.
It's clear that $\zeta^i(z_j) = I^i_j$.
The ζ^is form a basis for the dual space $\hat{V}$.
To extract the ith component of a dual vector α, let α act on z_i.
Let $T : V \to V$ be a linear transformation. Then transpose of T is the mapping from $\hat{V} \to \hat{V}$ such that: $\hat{T}(\alpha)(v) = \alpha(Tv)$. So, $\hat{T} \in L(\hat{V}, \hat{V})$.

Tensor Products

For each $v \in V, \alpha \in \hat{V}$, define the tensor product $v \otimes \alpha \in L(V, V)$ which acts on vectors $w \in V$ as follows: $(v \otimes \alpha)(w) = \alpha(w)v$.
If $(v \otimes \alpha)$ is regarded as a linear transformation, then it's transpose is given by $\beta \circ (v \otimes \alpha)$ for any $\beta \in \hat{V}$. And we have: $\beta \circ (v \otimes \alpha)(w) = \beta(v)\alpha(w)$, and so $\beta \circ (v \otimes \alpha) = \beta(v)\alpha$.
The full set of $z_i \otimes \zeta^j$ (with i and j taking all possible values) is a basis for $L(V, V)$.
Any linear transformation T from a vector space to itself can be expressed in the form $T^j_i z_j \otimes \zeta^i$, where the numbers T^j_i are called the components of T relative to the basis z_i.
$dim(L(V, V)) = dim(V)^2$.
$\zeta^i(T(z_j)) = \zeta^i(T^k_j z_k) = T^k_j I^i_k = T^i_j$.

Matrix

Square matrices form an algebra, i.e., a vector space that is closed under multiplication too.
If $\hat{T}$ is the transpose transformation of T, then:
$\zeta^i(T(z_j)) = z_j(\hat{T}\zeta^i)$.
$v \otimes \alpha w \otimes \beta = \alpha(w) v \otimes \beta$. Hence, $L(V, V)$ is actually an algebra, just like the vector space of $n \times n$ matrices.
The matrix of the product (composition) of linear transformations is the product of the matrices of those transformations.
Hence, the mapping from $L(V, V)$ to the space of all $n \times n$ matrices is an algebra isomorphism (also preserves the multiplication structure).

Change of Basis

Let $v \in V$ and let z be a basis for V. Then, z defines a vector in F^n by: $v^*_z = z^{-1}v$.
Let $\alpha \in \hat{V}$ and let z be a basis for V. Then, α defines a dual vector α^*_z in $\hat{F^n}$ by: $\alpha^*_z \equiv \alpha \circ z$.
Given any $T \in L(V, V)$, use a fixed basis z to define a linear map from F^n to itself by defining: $T^*_z \equiv z^{-1} \circ T \circ z$.
If you want to change the basis from z to y, you can define the linear transformation $P \equiv z^{-1} \circ y$. The matrix of P is given by expressing y_i in terms of the z_k and using the components to build the ith column. P is called the change of basis matrix.
Under a change of basis:

$$v^*_y = P^{-1}v^*_z,\ \alpha^*_y = \alpha^*_z P,\ T^*_y = P^{-1}T^*_z P$$

Chapter 4

A linear transformation from V to itself is called an **operator**.
A non-zero vector $v \in V$ is called an **eigenvector** of an operator T if its direction is not changed by T, i.e., $Tv = \lambda v$ for some scalar λ, which is called the eigenvalue associated with v.

- The set of all eigenvectors associated with λ is a subspace of V, and is called the eigenspace associated with λ.
- The eigenvectors and eigenvalues of a transformation on a real vector space need not be real.
- If two eigenvectors have different eigenvalues, then they cannot be parallel.
- If three eigenvectors have three different eigenvalues, then none of them can be expressed as a linear combination of the other two.
- In general, if any collection of eigenvectors have eigenvalues that are all different, then those eigenvectors are all linearly independent.

- The inverse is not true: it's possible to have a set of linearly independent vectors with the same eigenvalue.
- The number of distinct eigenvalues for an operator on a finite dimensional vector space cannot be larger than dim(V).

If v is an eigenvector with eigenvalue λ, any scalar multiple of v is also an eigenvector with eigenvalue λ.

Fundamental Theorem of Algebra: Any polynomial with complex coefficients can be completely factorized, i.e., it can be written in the form $a(x-a_1)(x-a_2)\cdots$, where the factorization is unique (up to the ordering).

Every operator on a finite-dimensional vector space over the complex numbers has at least one eigenvalue, a complex number (which may be real).

Upper Triangular Matrices

A matrix is upper-triangular if all entries below the diagonal are zero.

Triangularizability: Let T be an operator on a complex finite-dimensional vector space. Then there always exists a basis t for V such that, for all i, Tt_i is expressible as a linear combination of the t_j with $j \le i$.

Let T be any operator on a finite-dimensional vector space V, let z be any basis for V, and let T^*_z be the operator on F^n defined by T and z. Then for any z, T^*_z has the same eigenvalues as T.

So, we can find the eigenvalues of T by studying (any of) its matrices.

Square matrices A and B are said to be **similar** if there exists another (non-singular) matrix P such that $A = P^{-1}BP$.

Similar matrices have the same eigenvalues. If A satisfies some polynomial equation, B also satisfies that same polynomial equation.

The matrix of T relative to the special basis t (above) is upper-triangular. So, every matrix is similar to an upper-triangular matrix.

A complex operator is invertible if, and only if, every entry down the diagonal is non-zero when it is expressed as an upper-triangular matrix.

The diagonal entries of an upper-triangular matrix are its eigenvalues.

Eigentriangularizability: Let T be an operator on a complex finite-dimensional vector space. Then, there always exists a basis t for V such that, for all i, Tt_i is expressible as a linear combination of the t_j with all $j \le i$. Furthermore, the coefficient of t_i in this linear combination is an eigenvalue of T.

An operator is said to be **diagonalizable** if there is a basis with respect to which its matrix is diagonal. Clearly, this means that the chosen basis consists of all eigenvectors.

If $dim(V) = n$, and an operator on V has n distinct eigenvalues, then that operator is diagonalizable. But the inverse is not true: it's possible for an operator to have fewer than n distinct eigenvalues, and still be diagonalizable.

Jordan Canonical Form (JCF)

Every matrix has a Jordan Canonical Form.

Jordan Block: Let λ be any number. Then, a Jordan Block of size m is just an $m \times m$ matrix (here, $m = 4$) of the form:

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

Every Jordan Block (as a matrix itself) has only 1 eigenspace (corresponding to eigenvalue λ) which is one-dimensional. That is, each Jordan Block only contributes one (linearly independent from the rest) eigenvector.

So, the presence of Jordan Block of size ¿ 1 means you don't have enough eigenvectors to make a basis - and that's why you can't diagonalize it.

A Jordan basis is one such that the matrix of T consists of Jordan Blocks:

$$\begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix}$$

Every operator on a complex vector space has a Jordan basis. This matrix (not basis) is unique up to the ordering of the basis vectors.

The multiplicity of λ is the number of times the eigenvalue λ appears down the diagonal (in the JCF).

Let $\lambda_1, \lambda_2, \cdots$ be the eigenvalues of a linear operator T, and let m_i be the multiplicity of eigenvalue λ_i. Then the characteristic polynomial of T is the polynomial:

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

Clearly, for all i, $\chi_T(\lambda_i) = 0$.

Cayley-Hamilton Theorem:

$$\chi_T(T) = 0$$

The transformation (its matrix relative to any basis) satisfies its own characteristic equation.

If J is a Jordan Block of size m and having eigenvalue λ_i along its diagonal, then $(J - \lambda_i I)^m = 0$, i.e., it will map everything to the zero vector after applying it m times (hence, **nilpotent**).

We can also use the Cayley-Hamilton theorem to find the inverse of a matrix. For example, if M has 3 distinct eigenvalues 1, 2, 3. Then, the characteristic polynomial is:
$(x-1)(x-2)(x-3) = x^2 - 6x^2 + 11x - 6$. By CH, we have, $M^3 - 6M^2 + 11M + 6I = 0$. We can multiply both sides by M^{-1} (if it exists) and get $M^2 - 6M + 11 + 6M^{-1} = 0$ and solve for M^{-1}.

What the JCF Means

If M is a real matrix, then the complex conjugate of any eigenvalue is also an eigenvalue, i.e., eigenvalues come in complex-conjugate pairs.

We make use of the facts of complex numbers:

- $\overline{e^{i\theta} = \cos(\theta) + i\sin(\theta)}$
- $\overline{c+d} = \overline{c} + \overline{d}$, and $\overline{c \times d} = \overline{c} \times \overline{d}$ where $^-$ represents the conjugate operator

Also, we know that the eigenvalues of the rotation matrix are $e^{i\theta}$ and $e^{-i\theta}$.

Then, we can classify all linear transformations using their JCFs as follows:

- If a linear operator T on R^n has a real, diagonal JCF, then T has a geometric interpretation as either a sequence of stretches or reflections (if negative), or as a crushing (if zero) transformation.
- If a linear operator T on R^n has a diagonal, not necessarily real, JCF, then T has a geometric interpretation as either a sequence of stretches and reflections (when all eigenvalues are real), or as a crushing transformation (when one of the eigenvalue is zero), or as a sequence of stretches and rotations (when the eigenvalues are complex)
- If the JCF contains non-trivial Jordan blocks, but still with real numbers down the diagonal, then each Jordan block corresponds to a stretching (in all those dimensions by the same amount λ) and shearing (by the same amount, in each of the (i,i+1) planes).
- If the JCF contains non-trivial Jordan blocks with complex eigenvalues, it corresponds to a mixture of stretches, rotations, and shears.

Every linear operator T on R^n has a geometric interpretation as a combination of stretches, reflections, crushings, rotations, and shears.

Results from Tutorial

Tutorials 1-3

- Under a change of basis matrix $\mathbf{P}$, column vectors change from $\mathbf{c} \to \mathbf{P}^{-1}\mathbf{c}$, row vectors change from $\mathbf{r} \to \mathbf{rP}$, and matrices change from $\mathbf{M} \to \mathbf{P}^{-1}\mathbf{MP}$
- If U_1 and U_2 are subspaces of a finite-dimensional vector space V such that $V = U_1 \oplus U_2$, and ϕ is a vector space isomorphism from $V \to W$, then $W = \phi(U_1) \oplus \phi(U_2)$.

Tutorial 4

The mapping $K : V \to \hat{\hat{V}}$ defined by $K(v)(\alpha) = \alpha(v)$ is a natural isomorphism from $V \to \hat{\hat{V}}$.

Tutorial 5

- The tensor product gets turned into matrix multiplication when we represent things as matrices. e.g. $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (where the equality is by PoCS)
- Let $v \in V$ and $\alpha \in \hat{V}$, then the transpose of $v \otimes \alpha$ is given by $a \otimes K(v)$ where K is the canonical isomorphism from $V \to \hat{\hat{V}}$
- If $T : V \to W$, then the transpose of T, is $\hat{T} : \hat{W} \to \hat{V}$ (i.e., goes backwards) and defined by the usual definition of transpose, $\hat{T}(\alpha)(v) = \alpha(Tv)$ where $\alpha \in \hat{W}$ and $v \in V$.
- If $v, w \in V$ and $\alpha, \beta \in \hat{V}$ then:
$$(v \otimes \alpha) \circ (w \otimes \beta) = \alpha(w) v \otimes \beta$$
- $Tr(MN) = Tr(NM)$. In fact, a commutator is defined by $[M, N] = MN - NM$, and so, the trace of any commutator is zero. Also, any zero-trace matrix can be written as the commutator of some pair of matrices.

Tutorial 6

- Let z_i be a basis of a finite-dimensional vector space, and let P_i^j be a matrix. Then the vectors $y_i = P_i^j z_j$constitute a basis if, and only if, P_i^j is not singular. Basically, any bijection/isomorphism map one basis set to another basis set.
- Let η^j be the dual basis for y_i, and let ζ^j be the dual basis for z_i. Then:
$$\eta^j = (P^{-1})_i^j \zeta^i$$
Let $s_i \zeta^i$ be any dual vector. Show that the components, assembled into a row vector, change by multiplying by P_i^j on the right.
Finally, let $T_i^j z_j \otimes \zeta^i$ be any linear transformation from V to itself. Then, replacing z_i by y_i and ζ^i by η^i, we get our favorite $P^{-1}TP$ formula.
- A matrix is diagonalizable if, and only if, every eigenvalue has as eigenspace with dimensionality equal to its multiplicity (i.e., we have n linearly independent eigenvectors which can form a basis).