

Lecture 2. The Categories of Vector Spaces

PCMI Summer 2015 Undergraduate Lectures on Flag Varieties

Lecture 2. We discuss the categories of vector spaces and linear maps. Since vector spaces are always given in terms of a (scalar) field, we need to start with them.

A *field* is roughly a set of numbers with addition and multiplication operations that obey all the rules of arithmetic. You can get a long way by remembering that the rational numbers \mathbb{Q} and real and complex numbers \mathbb{R} and \mathbb{C} are fields, but the natural numbers \mathbb{N} (no negatives) and integers \mathbb{Z} (no reciprocals) are not. More precisely:

Definition 2.1. A set k with addition and multiplication operations $+$ and \cdot (the latter of which is suppressed in notation) is a *field* if:

$$a + (b + c) = (a + b) + c \quad \forall a, b, c \in k$$

$$a + b = b + a \quad \forall a, b \in k$$

$$\exists 0 \in k \text{ such that } 0 + a = a \quad \forall a \in k$$

Each a has an additive inverse $-a$ such that $a + (-a) = 0$

(*Subtraction* is addition of the additive inverse: $a - b := a + (-b)$)

$$a(bc) = (ab)c \quad \forall a, b, c \in k$$

$$ab = ba \quad \forall a, b \in k$$

$$\exists 1 \in k \text{ such that } 1a = a \quad \forall a \in k$$

Each $a \neq 0$ has a multiplicative inverse $1/a$ such that $a(1/a) = 1$

(*Division* is multiplication by the multiplicative inverse: $a/b := a(1/b)$)

$$a(b + c) = ab + ac \quad \forall a, b, c \in k$$

The field \mathbb{Q} of rational numbers is the smallest field containing \mathbb{Z} . The field \mathbb{R} of real numbers is the *completion* of \mathbb{Q} , consisting of all limits of convergent sequences of rational numbers. The field \mathbb{C} of complex numbers is the *algebraic closure* of \mathbb{R} , in which all polynomials (with real or complex coefficients) have a complex root.

There are many other interesting fields. There are fields with finitely many elements, fields of (rational) functions, and completions of \mathbb{Q} with respect to unusual (p -adic) notions of distance, as well as a wealth of *number fields* sitting between \mathbb{Q} and \mathbb{C} .

Example 2.1. For each prime p , the set of p elements:

$$\mathbb{F}_p = \{0, 1, \dots, p - 1\}$$

together with “mod p ” addition and multiplication is a field.

The *standard vector space* of dimension n over a scalar field k is:

$$k^n := k \times k \times \dots \times k$$

with componentwise addition of vectors and multiplication by scalars, i.e. if $\vec{v} = (v_1, \dots, v_n)$, $\vec{w} = (w_1, \dots, w_n) \in k^n$ and $c \in k$, then:

$$\vec{v} + \vec{w} = (v_1 + w_1, \dots, v_n + w_n) \quad \text{and} \quad c\vec{v} = (cv_1, \dots, cv_n)$$

Example 2.2. \mathbb{R}^2 is the Euclidean plane and \mathbb{R}^3 is Euclidean 3-space.

The key properties of k^n are gathered together in the following:

Definition 2.2. A *vector space over k* is a set V of vectors, equipped with an addition operation and an operation of multiplication by scalars that satisfies all the following properties:

- (a) $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w} \quad \forall \vec{u}, \vec{v}, \vec{w} \in V$
- (b) $\vec{u} + \vec{v} = \vec{v} + \vec{u} \quad \forall \vec{u}, \vec{v} \in V$
- (c) There is a zero vector $\vec{0} \in V$ satisfying $\vec{0} + \vec{v} = \vec{v} \quad \forall \vec{v} \in V$.
- (d) Every vector \vec{v} has an additive inverse $-\vec{v}$.
- (e) The multiplicative identity $1 \in k$ satisfies $1 \cdot \vec{v} = \vec{v} \quad \forall \vec{v} \in V$.
- (f) $c(c'\vec{v}) = (cc')\vec{v}$
- (g) $c(\vec{v} + \vec{w}) = c\vec{v} + c\vec{w}$
- (h) $(c + c')\vec{v} = c\vec{v} + c'\vec{v}$

Examples 2.3. (a) The standard vector spaces k^n .

(b) The zero vector space $\{\vec{0}\}$ (analogous to the empty set).

(c) Any nonempty subset $U \subset V$ of a vector space V (e.g. $V = k^n$) that is closed under addition and scalar multiplication is a vector space. Such a U is called a *subspace* of V . For example, the subspaces of Euclidean space \mathbb{R}^3 are $\{\vec{0}\}$ and the lines and planes through $\vec{0}$.

Definition 2.3. A map of vector spaces $f : V \rightarrow W$ over k is *linear* if:

- (i) $f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w}) \quad \forall \vec{v}, \vec{w} \in V$
- (ii) $f(c\vec{v}) = cf(\vec{v}) \quad \forall c \in k, \vec{v} \in V$

Exercise 2.1. (a) The identity map $\text{id}_V : V \rightarrow V$ is linear.

(b) The composition of a pair of linear maps is linear.

(c) The inverse of a linear map (if it exists) is linear.

Moment of Zen. The category $\mathcal{V}S_k$ of vector spaces over k is:

(a) The collection of vector spaces V over k , with (b) Linear maps.

Definition 2.4. The automorphisms $\text{Aut}(V)$ of V in the category $\mathcal{V}S_k$ make up the *general linear group* $\text{GL}(V)$.

Remark. The general linear group $\text{GL}(k^n)$ is usually written as $\text{GL}(n, k)$.

Linear maps of standard vector spaces have a convenient:

Matrix Notation: Let $e_1, \dots, e_n \in k^n$ be the *standard basis vectors*

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$$

and suppose $f : k^n \rightarrow V$ is a linear map. Then the list of images of basis vectors $f(e_1), \dots, f(e_n) \in V$ determines the linear map f since $\vec{v} = v_1 e_1 + \dots + v_n e_n$ (for unique scalars v_1, \dots, v_n) and:

$$f(\vec{v}) = v_1 f(e_1) + v_2 f(e_2) + \dots + v_n f(e_n) = \sum_{j=1}^n v_j f(e_j) \in V$$

If $V = k^m$, let $e'_1, \dots, e'_m \in k^m$ be the standard basis vectors and

$$f(e_j) = (a_{1,j}, \dots, a_{m,j}) = a_{1,j} e'_1 + \dots + a_{m,j} e'_m = \sum_{i=1}^m a_{i,j} e'_i$$

Then the $m \times n$ (m rows, n columns) array $A = (a_{i,j})$ is the *matrix notation* for $f : k^n \rightarrow k^m$, in which the j th column is the vector $f(e_j)$.

Proposition 2.1. If $f : k^n \rightarrow k^m$ and $g : k^m \rightarrow k^l$ are linear maps with matrix notation A and B , then the composition $h = g \circ f : k^n \rightarrow k^l$ has the $l \times n$ matrix notation given by the *matrix product*:

$$C = B \cdot A = \left(\sum_j b_{i,j} a_{j,r} \right) = (c_{i,r})$$

Proof. Letting e''_1, \dots, e''_l be a standard basis for k^l :

$$g(f(e_r)) = g\left(\sum_j a_{j,r} e'_j\right) = \sum_j a_{j,r} \left(\sum_i b_{i,j} e''_i\right) = \sum_i \left(\sum_j b_{i,j} a_{j,r}\right) e''_i$$

Definition 2.5. The *dot product* of two vectors in k^m is:

$$\vec{v} \cdot \vec{w} = (v_1, \dots, v_m) \cdot (w_1, \dots, w_m) = \sum_{j=1}^m v_j w_j$$

The dot product is **not** a multiplication of vectors, since it takes two vectors and produces a scalar. It is, however, commutative and (bi)-linear in the sense that $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$ and $(c\vec{v}) \cdot \vec{w} = c(\vec{v} \cdot \vec{w})$. (It doesn't even make sense to ask whether it is associative!) It is very useful in computations to notice that the entry $c_{i,r}$ in $C = B \cdot A$ is the dot product of the i th row $(b_{i,*})$ of B with the r th column $(a_{*,r})$ of A .

Example 2.4. Rotation of \mathbb{R}^2 by an angle θ counterclockwise is:

$$r_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

and a composition of rotations is the rotation by the sum of the angles:

$$\begin{aligned} r_{\psi+\theta} &= r_\psi \circ r_\theta = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\psi)\cos(\theta) - \sin(\psi)\sin(\theta) & -(\cos(\psi)\sin(\theta) + \sin(\psi)\cos(\theta)) \\ \sin(\psi)\cos(\theta) + \cos(\psi)\sin(\theta) & -\sin(\psi)\sin(\theta) + \cos(\psi)\cos(\theta) \end{bmatrix} \end{aligned}$$

which translates into the angle addition formulas for sin and cos.

Example 2.5. Multiplying the second column of r_θ by -1 gives:

$$f_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

which is the reflection across the line $y/x = \tan(\theta/2)$, and in particular:

$$f_\theta^2 = \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & 0 \\ 0 & \sin^2(\theta) + \cos^2(\theta) \end{bmatrix} = \text{id}_{\mathbb{R}^2}$$

is the Pythagorean theorem.

These two “circles” of linear maps are the symmetries of the circle.

Definition 2.6. The *determinant* of an $n \times n$ matrix $A = (a_{ij})$ is:

$$\det(A) = \sum_{\sigma \in \text{Perm}(n)} \text{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}$$

One way to think about the determinant. When $k = \mathbb{R}$, then $|\det(A)|$ is the “volume scaling factor” of the associated linear map, and if it is nonzero, then the sign of $\det(A)$ describes whether or not “orientation” is preserved by the linear map.

Another way to think about the determinant. We can think about an $n \times n$ matrix A as being the list of its column vectors. Then the determinant is a function that takes the list of column vectors as input and produces a scalar as output, analogous to the dot product (which takes two vectors as input and produces a scalar). Like the dot product, it is “multi”-linear, i.e. linear in each column vector, but unlike the dot product, the determinant function is anti-commutative in the sense that when two columns of a matrix are transposed, the determinant changes sign. These properties, together with $\det(\text{id}) = 1$, *completely determine* the determinant function.

In one case, the determinant is the sign of a single permutation:

Permutation Matrices. Consider $f_\tau : k^n \rightarrow k^n$ that permutes the standard basis vectors according to $\tau \in \text{Perm}(n)$:

$$f_\tau(e_i) = e_{\tau(i)}$$

The columns of the matrix P_τ for f_τ are the standard basis vectors permuted by τ , so that $p_{i,\tau^{-1}(i)} = 1$ for all i and all the other entries are zero (think about the τ^{-1} !). Evidently, these linear maps satisfy:

$$(f_{\tau'} \circ f_\tau)(e_i) = f_{\tau'}(e_{\tau(i)}) = e_{\tau'(\tau(i))} = f_{\tau' \circ \tau}(e_i)$$

so the assignment $\tau \mapsto f_\tau$ converts composition of permutations to composition of elements of the general linear group.

This is called the *natural representation* of $\text{Perm}(n)$. Notice that

$$\det(P_\tau) = \sum_{\sigma} \text{sgn}(\sigma) \prod_i p_{i,\sigma(i)} = \text{sgn}(\tau^{-1}) \prod_i p_{i,\tau^{-1}(i)} = \text{sgn}(\tau^{-1})$$

for the **single** permutation τ^{-1} , and $\text{sgn}(\tau) = 1/\text{sgn}(\tau^{-1}) = \text{sgn}(\tau^{-1})$.

Diagonal Matrices. The determinant of the diagonal $n \times n$ matrix:

$$D = \text{diag}\{d_1, \dots, d_n\} \text{ is the product } \prod_{i=1}^n d_i$$

More generally, the determinant of an upper or lower triangular matrix:

$$U = (u_{i,j}) \mid u_{i,j} = 0 \text{ for } i > j \text{ or } L = (l_{i,j}) \mid l_{i,j} = 0 \text{ for } i < j$$

is equal to the product of the terms on the diagonal (Exercise 2.2)

Transpose. The transpose of a matrix $A = (a_{i,j})$ is $A^T = (a_{j,i})$.

Remark. The transpose converts $m \times n$ matrices to $n \times m$ matrices!

The transpose is the matrix for the *dual* map, in the following sense.

Definition 2.7. The *dual* of the vector space V is the vector space:

$$V^\vee = \{\text{linear functions } x : V \rightarrow k\}$$

with addition of linear functions defined by $(x + y)(\vec{v}) = x(\vec{v}) + y(\vec{v})$ and multiplication by scalars defined by $(cx)(\vec{v}) = x(c\vec{v})$.

Duality. To each linear map $f : V \rightarrow W$ there is a *dual linear map*

$$f^\vee : W^\vee \rightarrow V^\vee \text{ defined by } (f^\vee \circ x)(\vec{v}) = x(f(\vec{v})).$$

If $f : k^n \rightarrow k^m$ has matrix A , let $x_1, \dots, x_n \in (k^n)^\vee$ be the *dual basis* defined by $x_i(e_i) = 1, x_i(e_j) = 0$ for $i \neq j$ and likewise for k^m . Then:

$$(f^\vee \circ x'_i)(e_j) = x'_i(f(e_j)) = x'_i\left(\sum a_{j,k} e_k\right) = a_{j,i}$$

so $f^\vee(x'_i) = \sum_j a_{j,i} x_j$ has the transpose matrix $A^T = (a_{j,i})!$

Dual maps compose in the reverse order:

$$(f \circ g)^\vee = g^\vee \circ f^\vee, \quad (A \cdot B)^T = B^T \cdot A^T$$

just as inverses do, but the dual is **not** (usually) the inverse!

For square matrices, it follows directly from the definition that:

$$\det(A^T) = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_i a_{\sigma(i),i} = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{\sigma^{-1}(i)} a_{i,\sigma^{-1}(i)} = \det(A)$$

once again since each $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Given a linear map $f : V \rightarrow W$ of vector spaces over k :

Definition 2.8. The *kernel* and *image* of f are the two subspaces:

$$\ker(f) = f^{-1}(\vec{0}) \subset V \text{ (a subspace of the domain of } f \text{) and}$$

$$\operatorname{im}(f) \subset W \text{ (a subspace of the range of } f \text{)}$$

Proposition 2.2. f is injective if and only if the kernel of f is $\{\vec{0}\}$.

Proof. f fails to be injective if and only if there are two vectors $\vec{v} \neq \vec{v}' \in V$ such that $f(\vec{v}) = f(\vec{v}')$, if and only if there is a nonzero vector $\vec{w} = \vec{v} - \vec{v}'$ in the kernel of f , since $f(\vec{w}) = f(\vec{v}) - f(\vec{v}') = \vec{0}$. \square

Associated to every subspace $U \subset V$, there is a **quotient** space.

Definition 2.9. The quotient vector space V/U is defined by:

(a) The vectors of V/U are the “cosets” $\vec{v} + U = \{\vec{v} + \vec{u} \mid \vec{u} \in U\}$.

(b) Addition of vectors of V/U and scalar multiplication is:

$$(\vec{v} + U) + (\vec{w} + U) = (\vec{v} + \vec{w}) + U \text{ and } c(\vec{v} + U) = c\vec{v} + U$$

Proposition 2.3. The natural map from V to the quotient space:

$$q : V \rightarrow V/U \text{ defined by } q(\vec{v}) = \vec{v} + U$$

is linear and surjective, and $U = \ker(q)$.

Proof. Surjectivity and linearity are built into the definition, and $U = \vec{0} + U$ is the origin in V/U . But $q^{-1}(U)$ consists of all vectors \vec{v} such that $\vec{v} + U = U$, i.e. the vectors of U ! \square

Corollary 2.1. A linear map $f : V \rightarrow W$ is surjective if and only if the dual map $f^\vee : W^\vee \rightarrow V^\vee$ is injective.

Proof. Suppose f is surjective and $x \in W^\vee$ is not the zero vector. Then $x : W \rightarrow k$ is not the zero map, so $f^\vee(x) = x \circ f : V \rightarrow W \rightarrow k$ is also not the zero map, i.e. $f^\vee(x) \neq \vec{0}$. In other words, $\ker(f^\vee) = \vec{0}$, so f^\vee is injective by Proposition 2.2.

On the other hand, if f fails to be surjective, let $U = \text{im}(f) \subset W$, and let $g : W \rightarrow W/U$ be the surjective map from Proposition 2.3. Then $W/U \neq \{\vec{0}\}$, and it follows that $(W/U)^\vee$ is not the zero space. Let $x \in (W/U)^\vee$ be a non-zero element. Then $g^\vee(x) \neq \vec{0}$, by the previous paragraph, but $f^\vee(g^\vee(x)) = \vec{0}$, since $g \circ f$ is the zero map, so f^\vee fails to be injective, as desired. \square

Now for the big theorem about determinants:

Theorem 2.1. Let A be the $n \times n$ matrix associated to $f : k^n \rightarrow k^n$.

(a) $\det(A) \neq 0$ if and only if f has an inverse map.

(b) The determinant is a multiplicative function:

$$\det(A \cdot B) = \det(A) \det(B)$$

Proof. “Row and column operations” on B give the following:

Basic Linear Algebra Fact. Every $n \times n$ matrix B is a product of:

(i) Transposition permutation matrices $P_{(i\ j)}$ for some pair $i \neq j$.

(ii) Diagonal matrices $D = \text{diag}\{d_1, \dots, d_n\}$ and

(iii) “Shearing” matrices S_{ci+j} with ones on the diagonal and all other entries are zero except for a single entry $s_{i,j} = c$.

When A is multiplied on the right by one of these matrices...

AP_τ permutes the columns of A according to τ .

AD multiplies the columns of A by the corresponding entries of D .

AS_{ci+j} adds c times the i th column of A to the j th column.

Part (b) of the Theorem now boils down to three computations:

$$\begin{aligned} \det(AP_\tau) &= \sum_{\sigma} \text{sgn}(\sigma) \prod_i a_{i,\sigma(\tau^{-1}(i))} \\ &= \text{sgn}(\tau) \sum_{\sigma} \text{sgn}(\sigma \circ \tau^{-1}) \prod_i a_{i,\sigma \circ \tau^{-1}(i)} = \det(A) \cdot \text{sgn}(\tau) = \det(A) \cdot \det(P_\tau) \end{aligned}$$

Remark. It follows from this computation that if two different columns of A are the same vector, then $A = AP_{(i\ j)}$, and so:

$$\det(A) = \det(AP_{(i\ j)}) = \det(A) \text{sgn}(i\ j) = -\det(A)$$

from which it follows that $\det(A) = 0$ for such matrices.

Next:

$$\det(A \cdot D) = \sum_{\sigma} \text{sgn}(\sigma) \prod_i d_{\sigma(i)} a_{i,\sigma(i)} = \left(\prod_i d_i \right) \cdot \det(A) = \det(A) \det(D)$$

Remark. It follows from this computation and the previous remark that if one column vector of A is equal to a *multiple* of another column vector of A , then $\det(A) = 0$.

Finally, suppose the j th column $f(e_j) = (a_{*,j})$ of a matrix A is the sum of two vectors: $\vec{v} + \vec{w} = (a_{*,j})$. Then:

$$\begin{aligned} \det(A) &= \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_i a_{i,\sigma(i)} = \sum_{\sigma} \operatorname{sgn}(\sigma) (v_{\sigma(j)} + w_{\sigma(j)}) \prod_{i \neq j} a_{i,\sigma(i)} \\ &= \det(A') + \det(A'') \end{aligned}$$

where A' and A'' are obtained by replacing the j th column of A with \vec{v} and with \vec{w} , respectively. This is the multi-linearity of the determinant. Applying this to the matrix AS_{ci+j} gives:

$$\det(AS_{ci+j}) = \det(A) + \det(A')$$

where A' is obtained from A by replacing the j th column of A with c times the i th column. But then A' is a matrix whose determinant is zero by the previous remark, so $\det(AS_{ci+j}) = \det(A)$. This completes the proof of (b).

Remark. It follows from this (and the previous remarks) that if one column vector of A is a *linear combination* of other column vectors, then $\det(A) = 0$. Indeed, suppose $f(e_j) = \sum_{i \neq j} c_i f(e_i)$. Then consider the matrix A' which agrees with A in all columns except that the j th column is the zero vector. By the previous remark, $\det(A') = 0$. But by a repeated multiplication by shearing matrices:

$$A = A' \cdot \prod_{i \neq j} S_{c_i, i+j} \quad \text{gives} \quad \det(A) = \det(A') = 0$$

Now, suppose f has an inverse map f^{-1} with matrix A^{-1} . Then

$$A \cdot A^{-1} = \operatorname{id} \quad \text{gives} \quad \det(A) \cdot \det(A^{-1}) = \det(\operatorname{id}) = 1$$

by Proposition 2.1 and (b), which was just proved. So $\det(A) \neq 0$.

Conversely, suppose f is not injective. Then there is some vector $\vec{w} = (w_1, \dots, w_n) \neq \vec{0}$, for which $f(\vec{w}) = w_1 f(e_1) + \dots + w_n f(e_n) = \vec{0}$ by Proposition 2.2. Suppose $w_j \neq 0$. Then:

$$f(e_j) = \sum_{i \neq j} \frac{w_i}{w_j} f(e_i)$$

exhibits $f(e_j)$ as a linear combination of other column vectors, so $\det(A) = 0$ by the final remark above. On the other hand, if f is not **surjective**, then the dual map f^\vee is not injective by Corollary 2.1, so $\det(A) = \det(A^T) = 0$, since A^T is the matrix for the dual map. \square

Exercises.

2.1. Show that:

- (a) The identity map $\text{id}_V : V \rightarrow V$ on a vector space is linear.
- (b) The composition of a pair of linear maps of vector spaces is linear.
- (c) The inverse of a linear map (if it exists) is linear.

2.2. Explain in two different ways why the determinants of upper and lower triangular matrices are the products of their diagonal terms:

- (a) Directly from the definition of the determinant.
- (b) By expressing an upper (lower) triangular matrix as the product of its diagonal and shearing matrices.

2.3. Theorem 2.1(a) only says that the linear map f associated to an $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$. It is also true that the associated map f is injective if and only if it is surjective, which is not implied by the theorem. Prove this using a more refined consequence of row and column operations:

Basic linear algebra fact: Any $n \times n$ matrix A is *diagonalizable* by invertible matrices B and C :

$$BAC = D$$

where B and C are products of permutation and shearing matrices.

2.4. Sketch an algorithm to come up with B and C as above, given A .

A vector space V is *finitely generated* if there is a surjective map:

$$f : k^N \rightarrow V \text{ for some } N$$

A vector space V *has dimension* n if there is an invertible map:

$$f : k^n \rightarrow V \text{ for some } n$$

in which case the n vectors $f(e_1), \dots, f(e_n)$ are called a *basis* of V .

2.5. (a) Prove that the dimension of a vector space is well-defined, i.e. prove that k^n is not isomorphic to k^m if $n \neq m$.

(b) Prove that if V is finitely generated, then V has finite dimension.

Definition. An $n \times n$ matrix A (or the associated linear map f) over \mathbb{C} is *semi-simple* if we may take $C = B^{-1}$ above, so that:

$$BAB^{-1} = D$$

Obviously, finding such a B is more challenging, but:

Definition. Let A be an $n \times n$ matrix. Then a nonzero vector $\vec{v} \in k^n$ is an *eigenvector* of A if:

$$A\vec{v} = \lambda\vec{v}$$

and the scalar λ is called the associated *eigenvalue*.

2.6. (a) Find a matrix $A \in \text{GL}(2, \mathbb{R})$ with no (real) eigenvectors.

(b) Show that **any** matrix $A \in \text{GL}(n, \mathbb{C})$ has an eigenvector.

Hint: Look for roots of the *characteristic polynomial*

$$\det(A - xI_n) = 0$$

where I_n is the identity matrix and use the fact that a polynomial with complex coefficients has (at least one) complex root.

(c) Find the eigenvalues and eigenvectors of rotation by θ .

(d) Show that the subset of k^n consisting of all eigenvectors of A with the same eigenvalue λ is a subspace of k^n . This is called the *eigenspace* associated to the eigenvalue λ .

2.7. Show that an $n \times n$ matrix A over \mathbb{C} is semi-simple if and only if \mathbb{C}^n has a basis consisting of eigenvectors. Show how to assemble matrices B and D above from a basis of eigenvectors of A .

2.8. Obviously a diagonal matrix is semi-simple. Show that:

(a) A shearing matrix is not semi-simple.

(b) A permutation matrix is semi-simple (always over \mathbb{C}).

Definition. The *trace* of $A = (a_{i,j})$ is the sum of its diagonal terms:

$$\text{tr}(A) = \sum_{i=1}^n a_{i,i}$$

2.9. If A and B are $n \times n$ matrices, prove that:

(a) $\text{tr}(AB) = \text{tr}(BA)$, and use this to conclude that:

(b) $\text{tr}(BAB^{-1}) = \text{tr}(A)$ if B is invertible.

Thus, in particular, if A is semi-simple, then:

$$\det(A) = \prod_{i=1}^n \lambda_i \text{ and } \text{tr}(A) = \sum_{i=1}^n \lambda_i$$

are the product and sum of the eigenvalues of A , as they appear in the diagonal matrix D (possibly with some appearing more than once).

(c) Reprove (b) by noticing that:

$$B(xI_n - A)B^{-1} = xI_n - BAB^{-1}$$

and therefore that the characteristic polynomials of A and BAB^{-1} are the same. But the characteristic polynomial of A is:

$$x^n - \operatorname{tr}(A)x^{n-1} + \cdots + (-1)^{n-1} \det(A)$$

so in particular, $\operatorname{tr}(A) = \operatorname{tr}(BAB^{-1})$.

2.10. Contemplate why Definition 2.9 is well-defined.

2.11. (a) Contemplate a throw-away remark from Corollary 2.1. The assertion was made that if a vector space V is not zero, then its dual V^\vee is not zero. How would you prove this? I.e. How would you find a non-zero element of the dual space?

(b) Find a “natural” linear map from V to its “double dual” $(V^\vee)^\vee$. Is such a map injective? (If so, then this shows that V^\vee is not zero.) For finitely generated vector spaces, argue that this map is invertible. For vector spaces that are not finitely generated, it might not be invertible. Why?

Zen master problems.

The collection of arrows $f : X \rightarrow Y$ in a category \mathcal{C} is denoted:

$$\operatorname{Hom}(X, Y)$$

In the category of vector spaces over k , this is commonly denoted by:

$$\operatorname{Hom}_k(V, W)$$

where V and W are vector spaces.

2.12. Show that each $\operatorname{Hom}_k(V, W)$ is a vector space over k , and that if $\dim(V) = n$ and $\dim(W) = m$, then $\dim(\operatorname{Hom}_k(V, W)) = nm$.

A **functor** $F : \mathcal{C} \rightarrow \mathcal{D}$ on categories consists of maps:

$$F : \operatorname{ob}(\mathcal{C}) \rightarrow \operatorname{ob}(\mathcal{D})$$

and, for every pair of objects X, Y from \mathcal{C} , a map:

$$F : \operatorname{Hom}(X, Y) \rightarrow \operatorname{Hom}(F(X), F(Y))$$

on the arrows, such that:

$$F(\operatorname{id}_X) = \operatorname{id}_{F(X)} \text{ and } F(f \circ g) = F(f) \circ F(g)$$

for all objects X and all composable arrows $g : X \rightarrow Y$ and $f : Y \rightarrow Z$.

2.13. Show that “forgetting the extra linear structure” is a functor from each category of vector spaces over k to the category of sets. But notice that “forgetting the complex structure” does **not** define a functor from vector spaces over \mathbb{C} to vector spaces over \mathbb{R} . Why not? After all, a vector space over \mathbb{C} is a vector space over \mathbb{R} !