

Lecture 5: Linear Algebra

This is the fifth week of the Mathematics Subject Test GRE prep course; here, we review the field of **linear algebra**!

1 Definitions and Concepts

Unlike the calculus we've been studying earlier, linear algebra is a field much more focused on its **definitions** than its applications. Accordingly, most of what you'll be asked to recall on a test are concepts and words, rather than any specific processes! We review as many of these concepts as we can here, though I recommend skimming through a linear algebra textbook for a more in-depth review.

1.1 Vector Spaces: Some Important Types

- **Vector space:** A **vector space** over a field F is a set V and a pair of operations $+: V \times V \rightarrow V$ and $\cdot: \mathbb{R} \times V \rightarrow V$, that are in a certain sense “well-behaved:” i.e. the addition operation is associative and commutative, there are additive identities and inverses, the addition and multiplication operations distribute over each other, the scalar multiplication is compatible with multiplication in F , and 1 is the multiplicative identity.¹

F will usually be \mathbb{R} on the GRE.

- Examples: $\mathbb{R}^n, \mathbb{C}^n, \mathbb{Q}^n$, the collection of all polynomials with coefficients from some field, the collection of all $n \times n$ matrices with entries from some field.

- **Subspace:** A subset S of a vector space V over a field F is called a **subspace** if it satisfies the following two properties: (1) for any $\mathbf{x}, \mathbf{y} \in S$ and $a, b \in F$, we have that $a\mathbf{x} + b\mathbf{y}$ is also an element of S , and (2) S is nonempty.
- **Span:** For a set S of vectors inside of some vector space V , the **span** of S is the subspace formed by taking all of the possible linear combinations of elements of S .
- **Row space:** For a $n \times k$ matrix M , the **row space** of M is the subspace of F^n spanned by the k rows of M .
- **Null space:** For a $n \times k$ matrix M , the **null space** of M is the following subspace:

$$\{\mathbf{x} \in F^n : M \cdot \mathbf{x} = \mathbf{0}\}.$$

- Useful Theorem: The orthogonal complement of the row space of a matrix M is the null space of M . Conversely, the orthogonal complement of the null space of a matrix M is the row space of M .

¹See [Wikipedia](#) if you want a precise description of these properties.

- **Eigenspace:** For any eigenvalue λ , we can define the **eigenspace** E_λ associated to λ as the space

$$E_\lambda =: \{\mathbf{v} \in V : A\mathbf{v} = \lambda\mathbf{v}\}.$$

1.2 Matrices: Some Important Types

- **Elementary Matrices:** There are three kinds of elementary matrices, which we draw below:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & \vdots & \dots & \vdots & \dots & 0 \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

The first matrix above multiplies a given row by λ , the second matrix switches two given rows, and the third matrix adds λ times one row to another row.

- **Reflection Matrices:** For a subspace U of \mathbb{R}^n , we can find a matrix corresponding to the map $\text{Refl}_U(\mathbf{x})$ by simply looking for eigenvectors. Specifically, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ form a basis for U and $\{\mathbf{w}_1, \dots, \mathbf{w}_{n-k}\}$ form a basis for U^\perp , note that the \mathbf{u}_i 's are all eigenvectors with eigenvalue 1, and the \mathbf{w}_i 's are all eigenvectors with eigenvalue -1 (because reflecting through U fixes the elements in U and flips the elements in U^\perp .) As a result, because we have n linearly independent eigenvectors, we can use our diagonalization construction (discussed later in these notes) EDE^{-1} to make a reflection matrix R .
- **Adjacency Matrices:** For a graph² G on the vertex set $V = \{1, 2, \dots, n\}$, we can define the **adjacency matrix** for G as the following $n \times n$ matrix:

$$A_G := \left\{ a_{ij} \mid \begin{array}{l} a_{ij} = 1 \quad \text{if the edge } (i, j) \text{ is in } E; \\ a_{ij} = 0 \quad \text{otherwise.} \end{array} \right\}$$

It bears noting that we can reverse this process: given a $n \times n$ matrix A_G , we can create a graph G by setting $V = \{1, \dots, n\}$ and $E = \{(i, j) : a_{ij} \neq 0\}$.

- Useful Theorem: In a graph G with adjacency matrix A_G , the number of paths from i to j of length m is the (i, j) -th entry in $(A_G)^m$.

- **Permutation:** A **permutation** σ of the list $(1, \dots, n)$ is simply some way to reorder this list into some other $(\sigma(1), \dots, \sigma(n))$. (If you prefer to think about functions, σ is simply a bijection from $\{1, \dots, n\}$ to $\{1, \dots, n\}$.)

²A **directed graph** $G = (V, E)$ consists of a set V , which we call the set of *vertices* for G and a set $E \subset V^2$, made of ordered pairs of vertices, which we call the set of *edges* for G .

Given any permutation σ of $(1, \dots, n)$, the **permutation matrix** P_σ is simply the $n \times n$ matrix whose i -th column is given by $e_{\sigma(i)}$. In other words,

$$P_\sigma = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ e_{\sigma(1)} & e_{\sigma(2)} & \dots & e_{\sigma(n)} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

1.3 Various Vector/Vector Space Properties

- **Dimension:** The **dimension** of a space V is the number of elements in a basis for V .
- **Rank:** The **rank** of a matrix is the dimension of its row space.
- **The rank-nullity theorem.** The rank-nullity is the following result:

Theorem. Let U, V be a pair of finite-dimensional vector spaces, and let $T : U \rightarrow V$ be a linear map. Then the following equation holds:

$$\text{dimension}(\text{null}(T)) + \text{dimension}(\text{range}(T)) = \text{dimension}(U).$$

- **Orthogonality:** Two vectors \mathbf{u}, \mathbf{v} are called **orthogonal** iff their inner product is 0; i.e. if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
 - Useful Theorem: If we have a basis B for some space V , the **Gram-Schmidt process** will transform B into an orthogonal basis U for V – i.e. a basis for V that's made of vectors that are all orthogonal to each other. See [my notes](#) for an in-depth description of this process.
- **Linear independence/dependence:** A collection $v_1 \dots v_k$ of vectors is called **linearly dependent** iff there are k constants $a_1 \dots a_k$, not all identically 0, such that $\sum_{i=1}^k a_i v_i = 0$. They are called **linearly independent** if no such collection exists.
 - Useful Theorem: A collection of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly dependent iff the matrix formed by taking the \mathbf{v}_i 's as its rows has a zero row in its reduced row-echelon form.
 - Equivalently, a collection of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly dependent iff the determinant of the matrix formed by taking the \mathbf{v}_i 's as its rows is zero.
- **Basis:** A **basis** for a space V is a collection of vectors B , contained within V , that is linearly independent and spans the entire space V . A basis is called **orthogonal** iff all of its elements are orthogonal to each other; it is called **orthonormal** iff all of its elements are orthogonal to each other and furthermore all have length 1.
- **Eigenvector/eigenvalue:** For a matrix A , vector \mathbf{x} , and scalar λ , we say that λ is an **eigenvalue** for A and \mathbf{x} is a **eigenvector** for A if and only if $A\mathbf{x} = \lambda\mathbf{x}$.

- **Algebraic multiplicity:** The **algebraic multiplicity** of an eigenvalue μ is the number of times it shows up as a root of A 's characteristic polynomial. I.e. if $p_A(\lambda) = (\lambda - \pi)^2$, π would have algebraic multiplicity 2.
- **Geometric multiplicity:** The **geometric multiplicity** of an eigenvalue μ is the dimension of the eigenspace associated to μ .
- Useful Theorem: The algebraic multiplicity of an eigenvalue is always greater than or equal to the geometric multiplicity of that eigenvalue.
- Useful Theorem: A matrix is diagonalizable iff every eigenvalue has its algebraic multiplicity equal to its geometric multiplicity. (If you want it to be diagonalizable via real-valued matrices, you should also insist that the matrix and all of its eigenvalues are real.)
- **Dominant eigenvalue:** The **dominant eigenvalue:** is the largest eigenvalue of a matrix.

1.4 Various Matrix Properties

- **Symmetric:** A matrix is called **symmetric** iff $A^T = A$.
 - Useful Theorem: $(AB)^T = B^T A^T$.
- **Singular/Nonsingular:** A $n \times n$ matrix is called **singular** iff it has $\text{rank} < n$, and is called **nonsingular** iff it has $\text{rank} = n$.
 - Useful Theorem: A matrix is nonsingular if and only if it has an inverse.
 - Useful Theorem: A matrix is nonsingular if and only if its determinant is nonzero.
- **Orthogonal:** A $n \times n$ matrix U is called **orthogonal** iff all of its columns are of length 1 and orthogonal to each other. Equivalently, U is orthogonal iff $U^T = U^{-1}$; i.e. $U^T U = U U^T = I$.
 - Useful Theorem: Any $n \times n$ orthogonal matrix can be written as the product of no more than $n - 1$ reflections. (Specifically, no more than $n - 1$ reflections through spaces of dimension $n - 1$, which we call hyperplanes.)
- **Regular:** A matrix A is called regular if $a_{ij} > 0$, for every entry a_{ij} in A . We will often write $A > 0$ to denote this.
- **Nonnegative:** A matrix is called **nonnegative** if and only if all of its entries are ≥ 0 .
 - Useful Theorem: Suppose that A is a nonnegative matrix and λ is the maximum of the absolute values of A 's eigenvalues. Then λ is itself an eigenvalue, and there is a vector of nonnegative numbers that is an eigenvector for λ .
 - Perron-Frobenius: If A is a nonnegative matrix such that $A^m > 0$ for some value of m , then the nonnegative eigenvector above is unique, up to scalar multiplication.

- If λ is an eigenvector of a nonnegative matrix A that corresponds to a nonnegative eigenvector, then λ is at least the minimum of the row sums, and at most the maximum of the row sums; similarly, λ is at least the minimum of the column sums, and at most the maximum of the column sums.
- **Similarity.** Two matrices A, B are called **similar** if there is some matrix U such that $A = UBU^{-1}$. If we want to specify what U is, we can specifically state that A and B are similar **via** U .
- **Diagonalizable:** A **diagonalization** of a matrix A is an orthogonal matrix E and a diagonal matrix D such that $A = EDE^{-1}$.
 - Useful Theorem: A matrix A is diagonalizable if and only if it has n linearly independent eigenvectors $\mathbf{e}_1, \dots, \mathbf{e}_n$. In this case, if $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues to the \mathbf{e}_i 's, we can actually give the explicit diagonalization of A as

$$\left(\begin{array}{c|c|c|c} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ | & | & & | \end{array} \right) \cdot \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} \cdot \left(\begin{array}{c|c|c|c} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ | & | & & | \end{array} \right)^{-1}$$

- Suppose that A is diagonalized as EDE^{-1} . Then we can write the n -th power of A as ED^nE^{-1} . As well, if all of the entries along the diagonal of D have k -th roots, we can give a k -th root of A as the product $ED^{1/k}E^{-1}$.
- **Positive-definite/positive-semidefinite:** A matrix A is called **positive-definite** iff for any nonzero vector \mathbf{x} , we have $\mathbf{x}^T \cdot A \cdot \mathbf{x} > 0$. Similarly, it is called **positive-semidefinite** iff for any nonzero vector \mathbf{x} , we have $\mathbf{x}^T \cdot A \cdot \mathbf{x} \geq 0$.
 - Useful Theorem: A matrix is positive-definite iff all of its eigenvalues are positive; similarly, a matrix is positive-semidefinite iff all of its eigenvalues are nonnegative.
- **Probability:** A $n \times n$ matrix P is called a **probability matrix** if and only if the following two properties are satisfied:
 - $P \geq 0$; in other words, $p_{ij} \geq 0$ for every entry p_{ij} of P .
 - The column sums of P are all 1; in other words, $\sum_{i=1}^n p_{ij} = 1$, for every j .
 - Useful Theorem: Every probability matrix has a stable vector.
 - Useful Theorem: If P is a probability matrix such that there is a value of m where $P^m > 0$, then there is only one stable vector \mathbf{v} for P . Furthermore, for very large values of m , P^m 's columns all converge to \mathbf{v} . This theorem also holds in the case where the graph represented by P is **strongly connected**³, even if P^m is never > 0 .

³A graph is **strongly connected** iff it's possible to get from any node to any other node via edges in the graph.

- Useful Theorem 3: If we have a probability matrix P representing some finite system with n states $\{1, \dots, n\}$, then the probability of starting in state j and ending in state i in precisely m steps is the (i, j) -th entry in P^m .
- **Polar decomposition:** For a nonsingular $n \times n$ matrix A , a **polar decomposition** of A is a pair of matrices Q, S such that Q is an orthogonal matrix and S is a positive-definite symmetric matrix.
- **Singular Value Decomposition:** For a $m \times n$ matrix A , a **singular value decomposition** of A is a $n \times n$ orthogonal matrix V , a $m \times n$ matrix D such that $d_{ij} \neq 0$ only when $i = j$, and a $m \times m$ orthogonal matrix U such that $A = UDV^T$.
 - Useful Theorem: If A has a singular value decomposition given by UDV^T , then A 's Moore-Penrose pseudoinverse A^+ is given by the product VD^+U^T , where D^+ is the $n \times m$ matrix formed by taking D 's transpose and replacing all of its nonzero entries with their reciprocals.
 - Useful Theorem: If A is a $n \times n$ matrix with SVD UDV^T , then the minimum value of $\|A\mathbf{x}\|/\|\mathbf{x}\|$ can be found by plugging in \mathbf{v}_i , where \mathbf{v}_i is the column of V corresponding to the smallest value d_{ii} on D 's diagonal.
- **Moore-Penrose pseudoinverse:** For a matrix A , we say that A^+ is the **pseudoinverse** of A iff the following four properties hold: (1) $AA^+A = A$, (2) $A^+AA^+ = A^+$, (3) AA^+ is symmetric, and (4) A^+A is also symmetric.
 - Useful Theorem: The least-squares best-fit solutions to the system $A\mathbf{x} = \mathbf{b}$ are given by vectors of the form

$$A^+ \cdot \mathbf{b} + (I - A^+A)\mathbf{w},$$

where we let \mathbf{w} be any vector. Furthermore, if there is a solution to $A\mathbf{x} = \mathbf{b}$, then $A^+ \cdot \mathbf{b}$ is a solution of minimum length.

- **The spectral theorem.** Suppose that A is a $n \times n$ real symmetric matrix (i.e. don't make any assumptions about what U is like we did above.) Then in A 's Schur decomposition URU^{-1} , R is a diagonal real-valued matrix! Furthermore, we can insure in our construction of U that it is a real-valued orthogonal matrix.
- **QR-decomposition.** A **QR-decomposition** of an $n \times n$ matrix A is an orthogonal matrix Q and an upper-triangular⁴ matrix R , such that

$$A = QR.$$

Every invertible matrix has a QR-decomposition, where R is invertible.

⁴A matrix is called **upper-triangular** if all of its entries below the main diagonal are 0. For example, $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ is upper-triangular.

- **Jordan block.** A block B_i of some block-diagonal matrix is called a **Jordan block** if it is in the form

$$\begin{bmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}.$$

In other words, there is some value λ such that B_i is a matrix with λ on its main diagonal, 1's in the cells directly above this diagonal, and 0's elsewhere.

- **Jordan canonical/normal form.** Suppose that A is similar to an $n \times n$ block-diagonal matrix B in which all of its blocks are Jordan blocks; in other words, that $A = UBU^{-1}$, for some invertible U . We say that any such matrix A has been written in **Jordan canonical form**.

Any $n \times n$ matrix A can be written in Jordan canonical form.

1.5 Operations on Vectors and Vector Spaces

- **Dot product:** For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we define the **dot product** $\mathbf{x} \cdot \mathbf{y}$ as the sum $\sum_{i=1}^n x_i y_i$.
- **Inner product:** For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we define the inner product $\langle \mathbf{x}, \mathbf{y} \rangle$ of \mathbf{x} and \mathbf{y} as their dot product, $\mathbf{x} \cdot \mathbf{y}$.
 - Useful Observation: Often, it's quite handy to work with the transpose of certain vectors. So, remember: when you're taking the inner product or dot product of two vectors, taking the transpose of either vector doesn't change the results! I.e. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}^T, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y}^T \rangle = \langle \mathbf{x}^T, \mathbf{y}^T \rangle$. We use this a *lot* in proofs and applications where there are symmetric or orthogonal matrices running about.
- **Magnitude:** The magnitude of a vector \mathbf{x} is the square root of its inner product with itself: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. This denotes the distance of this vector from the origin.
- **Distance:** The distance of two vectors \mathbf{x}, \mathbf{y} from each other is the square root of the inner product of the difference of these two vectors: $\|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$.
- **Projection, onto a vector:** For two vectors \mathbf{u}, \mathbf{v} , we define the projection of \mathbf{v} onto \mathbf{u} as the following vector:

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) := \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \cdot \mathbf{u}.$$

- **Projection, onto a space:** Suppose that U is a subspace with orthogonal basis $\{b_1, \dots, b_n\}$, and \mathbf{x} is some vector. Then, we can define the **orthogonal projection**

of \mathbf{x} onto U as the following vector in U :

$$\text{proj}_U(\mathbf{x}) = \sum_{i=1}^n \text{proj}_{\mathbf{b}_i}(\mathbf{x}).$$

– Useful Theorem: This vector is the closest vector in U to \mathbf{x} .

- **Orthogonal complement:** For a subspace S of a vector space V , we define the **orthogonal complement** S^\perp as the following set:

$$S^\perp = \{v \in V : \langle v, s \rangle = 0, \forall s \in S\}.$$

- **Isometry:** A **isometry** is a map $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that preserves distances: i.e. for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

$$\|\mathbf{x} - \mathbf{y}\| = \|f(\mathbf{x}) - f(\mathbf{y})\|$$

- **Reflection:** For a subspace U of \mathbb{R}^n , we define the **reflection map through U** as the function

$$\text{Refl}_U(\mathbf{x}) = \mathbf{x} - 2 \cdot \text{proj}_{U^\perp}(\mathbf{x})$$

1.6 Operations on Matrices

- **Transpose:** For a $m \times n$ matrix A , the transpose A^T is the $n \times m$ matrix defined by setting its (i, j) -th cell as a_{ji} , for every cell (i, j) .

- **Determinant** For a $n \times n$ matrix A , we define

$$\det(A) = \sum_{i=1}^n (-1)^{i-1} a_{1i} \cdot \det(A_{1i}).$$

– Properties of the Determinant:

- * Multiplying one of the rows of a matrix by some constant λ multiplies that matrix's determinant by λ ; switching two rows in a matrix multiplies the that matrix's determinant by -1 ; adding a multiple of one row to another in a matrix does not change its determinant.
- * $\det(A^T) = \det(A)$.
- * $\det(AB) = \det(A) \det(B)$.
- * The determinant of the matrix A is the volume of the parallelepiped spanned by the columns of A (up to a factor of ± 1 , which is how we determine if the map is orientation-preserving or -reversing.)

– Useful Theorem: The determinant of a matrix A is nonzero if and only if A is invertible.

- **Trace:** The trace of a $n \times n$ matrix A is the sum of the entries on A 's diagonal.

– Useful Theorem: The trace of a matrix is equal to the sum of its eigenvalues.

- **Characteristic polynomial:** The characteristic polynomial of a matrix A is the polynomial $p_A(\lambda) = \det(\lambda I - A)$, where λ is the variable.

– x is a root of $p_A(\lambda)$ iff x is an eigenvalue for A .

2 Example problems

We work some sample problems here, to illustrate some of the ideas.

Question. Suppose that A is an $n \times n$ matrix such that A^3 is the all-zeroes matrix, i.e. the $n \times n$ matrix in which every entry is 0.

1. Think of A as a linear map from $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Can the range of A be equal to \mathbb{R}^n ?
2. Can you find an example of such a matrix A , such that A and A^2 are not themselves all-zeroes matrices?

Proof. For an example of such a matrix, consider

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We can easily check that

$$A^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$A^3 = A^2 \cdot A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If you want a $n \times n$ example of such a matrix, simply add additional rows/columns of zeroes to the left and bottom of A .

In general, we claim that the range of any such matrix A cannot be \mathbb{R}^n . To see why, simply notice that if A is a matrix with range equal to its domain, then A must be invertible; consequently, for any natural number k , A^k must also be invertible, with inverse given by $(A^{-1})^k$. Therefore A^k would have range \mathbb{R}^n (as it is invertible, and thus has $\dim(\text{nullspace}) = 0$;) and therefore in particular we could not have $A^k =$ the all-zeroes matrix for any k , as this has $\dim(\text{nullspace}) = n$. \square

Question. Take any $n \times n$ matrix M .

1. Take any $k > 0 \in \mathbb{N}$, and think of M, M^k as a pair of linear maps from $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Prove that

$$\text{nullspc}(M^k) \supseteq \text{nullspc}(M).$$

2. Suppose that $\det(M) = 0$. Prove that $\det(M^k) = 0$ as well.

Proof. The first claim here is not hard to establish. Take any vector $\vec{v} \in \text{nullspc}(M)$. By definition, we know that $M\vec{v} = \vec{0}$; therefore, we can conclude that $M^k\vec{v} = M^{k-1} \cdot M\vec{v} = M^{k-1}\vec{0} = \vec{0}$ as well, and thus that $\vec{v} \in \text{nullspc}(M^k)$.

As a side note, our earlier problem proves that inequality is possible (as the nullspaces of A, A^3 were distinct;) it is also not hard to see that equality is possible (let M be the all-zeroes matrix!) and thus that this is the strongest statement we can make.

For the second part of our claim: we could simply use the multiplicative property of the determinant (which tells us that $\det(M^k) = \det(M) \cdot \dots \cdot \det(M) = 0 \cdot \dots \cdot 0 = 0$), or we could use the first part of this question to note that because

- $\det(M) = 0$ if and only if $\dim(\text{nullspc}(M)) \neq 0$, and
- $\text{nullspc}(M^k) \supseteq \text{nullspc}(M)$, then $\dim(\text{nullspc}(M)) \leq \dim(\text{nullspc}(M^k))$,
- then we can conclude that if $\det(M) = 0$ then $\dim(\text{nullspc}(M^k)) \neq 0$, and thus that
- $\det(M^k) = 0$.

□

Question. Create a 4×4 matrix A with the following properties:

- No entry of A is 0.
- A has 1,2,and 3 as eigenvalues.

Proof. If we ignore our “no zero entries” condition, this is not too hard; the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

satisfies our eigenvalue properties, as $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ are eigenvectors for these three eigenvalues 1, 2, 3

Now, we can use the fact that eigenvalues are invariant under similarity; that is, if A is a matrix and B is an invertible matrix, then A and BAB^{-1} have the same eigenvalues! (This is because if \vec{v} is an eigenvector for A , then $B\vec{v}$ is an eigenvector for BAB^{-1} , with the same eigenvalue.)

So we can try simply multiplying A on the left and right by appropriate B, B^{-1} 's, and hope we get something without zeroes! In particular, let's use some matrices whose inverses we know: elementary matrices! Recall that

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow B^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

because the first map (when processed as $B \cdot (\text{matrix})$) corresponds to the Gaussian elimination move of “add one copy of row two to row three,” and the second is just “add -1 copies of row two to row three.”

Therefore

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix};$$

by using more of these elementary matrices, we can actually get

$$\begin{aligned} & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \\ & \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 2 \\ -6 & 4 & 4 \\ 0 & 0 & 3 \end{bmatrix}; \\ & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 2 \\ -6 & 4 & 4 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ -2 & 4 & 4 \\ 2 & -1 & 1 \end{bmatrix}. \end{aligned}$$

This is a matrix that has no nonzero entries, and by construction is similar to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$;

so we've answered our problem! \square

Question. Suppose that A is a $n \times n$ matrix with the following two properties:

- A^n is the all-zeroes matrix.
- There is exactly one nonzero vector \vec{v} , up to scalar multiples, that is an eigenvector of A . (In other words, the only eigenvectors for A are vectors of the form $c \cdot \vec{v}$.)

Find the Jordan normal form of A .

Proof. Take any eigenvector \vec{v} for A ; then $A\vec{v} = \lambda\vec{v}$. Consequently, $A^n\vec{v} = A^{n-1}\lambda\vec{v} = A^{n-2}\lambda^2\vec{v} = \dots = \lambda^n\vec{v}$. Because A^n is the all-zeroes matrix, we can also observe that $A^n\vec{v} = \vec{0}$, for any vector \vec{v} ; consequently, we have proven that the only possible eigenvalue for A is 0.

Our second bullet point is the claim that the dimension of the eigenspace for this only eigenvalue is 1. Consequently, if we look at our matrix's Jordan normal form, we know that

- The diagonals are all zeroes, as 0 is the only eigenvalue, and eigenvalues go on the diagonal of a Jordan normal form.
- There is only one block, as there is only one dimension of eigenvectors.

Therefore, we have that the Jordan normal form here is just zeroes on the diagonal, ones directly about the diagonal, and zeroes elsewhere: i.e.

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

\square

Question. Suppose that A is a real-valued symmetric $n \times n$ matrix with the following two properties:

- All of A 's entries are either 0 or 1.
- The all-1's vector is an eigenvector of A , with eigenvalue 10.

1. How many 1's are there in each row of A ?
2. Suppose that $\lambda \neq 10$ is another eigenvalue of A . Prove that $\lambda \leq 10$.

Proof. Let $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$. Notice that by simply multiplying it out, $A \cdot (1, 1, \dots, 1)$ is the vector

$$\left(\sum_{i=1}^n a_{1i}, \sum_{i=1}^n a_{2i}, \dots, \sum_{i=1}^n a_{ni} \right).$$

If this is equal to $(10, 10, \dots, 10)$, then there are ten 1's in each row of A , as claimed.

Furthermore, suppose that we have any eigenvalue λ other than 10 for this matrix A . Let \vec{v} be the eigenvector for this eigenvalue, and v_k be the largest component of this eigenvector. Then, again by definition, we have

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}v_i \\ \vdots \\ \sum_{i=1}^n a_{ni}v_i \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

In particular, if we look at the v_k coordinate, we have

$$\sum_{i=1}^n a_{ki}v_i = \lambda v_k;$$

but if we use the fact that v_k is the "biggest" (i.e. $v_k \geq v_j, \forall j$), we can see that

$$\sum_{i=1}^n a_{ki}v_i \leq \sum_{i=1}^n a_{ki}v_k \leq 10v_k,$$

because there are at most ten one-entries in the k -th row (and the rest are zeroes.)

But this means that $\lambda v_k \leq 10v_k$; i.e. $\lambda \leq 10$, as claimed. □