Math 1b: The Final Review

1 Definitions, Theorems, and Tools

1.1 Vector Spaces: Some Important Types

- **Vector space**: A **vector space** over $\mathbb{R}$ is a set V and a pair of operations $+ : V \times V \to V$ and $\cdot : \mathbb{R} \times V \to V$, that are in a certain sense "well-behaved:" i.e. the addition operation is associative and commutative, there are additive identites and inverses, the addition and multiplication operations distribute over each other, the scalar multiplication is compatible with multiplication in $\mathbb{R}$, and 1 is the multiplicative identity.[1]
  - Example 1: $\mathbb{R}^n$.
  - Example 2: $\mathbb{C}^n$.
- **Subspace**: A subset S of a vector space V is called a **subspace** if it satisfies the following two properties: (1) for any $\mathbf{x}, \mathbf{y} \in S$ and $a, b \in \mathbb{R}$, 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 $\mathbb{R}^n$ 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:

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

  - 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.
  - Useful Theorem: If M is a $n \times n$ matrix, then $(\text{rowspace}(M) + \text{nullspc}(M)) = n$.
  - The handout on row and null spaces on the class website and my own handout here both explain in great detail how to find these objects.
- **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] See Wikipedia if you want a precise description of these properties.

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 $\mathrm{Refl}_U(\mathbf{x})$ by simply looking for eigenvectors. Specfically, 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 graph[2] 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:

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

  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 $(A_G)^m$.

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.

- **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 the class website or my notes for an in-depth description of this process.

- **Linear indepdendence/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.

[2] A **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.

- 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.

- **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 the geometric multiplicty 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: $(AB)^T = B^T A^T$.
  - Useful Theorem: Real symmetric matrices are diagonalizable by an orthogonal matrix. In other words, if A is a real symmetric matrix, then there is an orthogonal matrix E made out of A's eigenvectors and a diagonal matrix D made out of A's eigenvalues such that $A = EDE^T$. (Because E is orthogonal, we are justified in writing E^T in place of E^{-1}, as for orthogonal matrices these are the same thing.)

    The process for finding this diagonalization is the exact same as the one we use for normal diagonalization (find the eigenvalues, find their eigenspaces, and find a basis for each eigenspace; then, take all of those vectors and use them as E's columns, and put the eigenvectors in D), except we find orthonormal bases for each of the eigenspaces, instead of just normal bases.

- **Singular/Nonsingular**: A $n \times n$ matrix is called **singular** iff it has rank $< n$, and is called **nonsingular** iff it has 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. $U^T U = UU^T = I$.
  - 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.
- **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 $\mathbf{e}_1, \ldots \mathbf{e}_n$. In this case, if $\lambda_1, \ldots \lambda_n$ are the corresponding eigenvalues to the $\mathbf{e}_i$'s, we can actually give the explicit diagonalization of A as

$$\begin{pmatrix} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \\ | & | & & | \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \\ | & | & & | \end{pmatrix}^{-1}$$

  - 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.

[3] A 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 $\{1, \ldots n\}$, 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.

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 $f : \mathbb{R}^n \to \mathbb{R}^n$ 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.
    * $\det(A^T) = \det(A)$.
    * $\det(AB) = \det(A)\det(B)$.
  - 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.

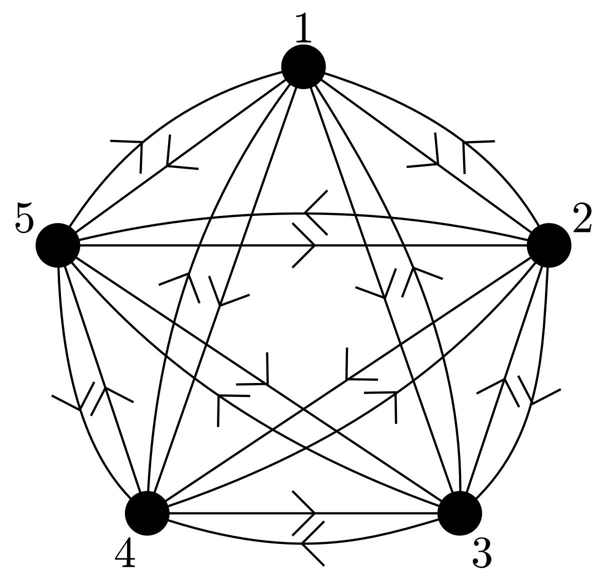
  - x is a root of $p_A(\lambda)$ iff x is an eigenvalue for A.

2 Examples

In this section, we work a series of interconnected problems, that (I think!) should illustrate most of the concepts above. (The only things I can think of that we don't really talk about here that occur above are reflections, which will show up in the week 10 rec notes I'll hopefully have up later today, and elementary matrices, which I covered in these notes pretty convincingly.

So: let's start!

Question 1 *Suppose that G is the following graph:*



How many paths are there from 1 to itself of length 712?

Solution. How can we solve this question? Well: recall the following theorem, which is somewhere in the above list of definitions and results:

Theorem 2 *If G is a graph with vertex set $\{1, \ldots n\}$, then the number of paths of length m from i to j in G is the (i,j)-th entry in $(A_G)^m$, where A_G is the adjacency matrix of G.*

Recall that the adjacency matrix for a graph is just the $n \times n$ matrix that has $a_{ij} = 1$ iff there is an edge from i to j, and 0 otherwise. This means that, in specific, the adjacency matrix for *our* graph G is the following matrix:

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

as each vertex has an edge from itself to every other vertex *other* than itself.

Now, we want to raise this matrix to a very large power. Direct calculations – even if you have access to Mathematica – are going to take forever and be kinda awful. What else can we do?

Well: if this matrix is **diagonalizable**, we can raise it to large powers with absolutely no effort at all! (This is because if we can write $A_G = EDE^{-1}$, then $(A_G)^{712} = (EDE^{-1}) \cdot (EDE^{-1}) \cdot \ldots \cdot (EDE^{-1})$, which if we cancel out all of the intermediate EE^{-1} terms becomes $ED^{712}E^{-1}$. Because raising a diagonal matrix to a large power is just raising all of its entries to that large power, we've reduced our massive number of calculations to simply multiplying three matrices!)

Is this matrix diagonalizable? In fact, it is – and we can tell this before we do any calculations at all, because it's a **symmetric** matrix! This means that it's actually diagonalizable by an **orthogonal matrix** – i.e. there is an orthogonal matrix E (i.e. E such that $E^{-1} = E^T$) and diagonal matrix D such that $A_G = EDE^T$. So, for practice's sake, let's diagonalize A_G with an orthogonal matrix.

How do we do this? The recipe we outlined for finding such "orthogonal" diagonalizations goes roughly as follows:

1. Find all of A_G's eigenvalues.

2. Using these eigenvalues, find the eigenspaces corresponding to each of these eigenvalues.

3. Now, find an **orthonormal basis** for each of these spaces – i.e. find a basis for each space made out of vectors that are all orthogonal to each other and all of length 1.

4. Take all of the vectors you got in step (3) above, and use them as the columns of a matrix E; by definition, this is an orthogonal matrix. Then let D be the diagonal matrix full of A_G's eigenvalues, placed so that the eigenvalue at (i, i) corresponds to the i-th column in E. Then $A_G = EDE^T$! and we're done.

Straightforward enough. Let's calculate!

To find A_G's eigenvalues, there are two ways. The first way would be to simply calculate $\det(\lambda I - A)$, A_G's characteristic polynomial, and find its roots; this will work, but for a 5×5 matrix seems kind of tedious.

The second way – and it bears noting that this is a completely valid philosophy to use when doing maths! – is **to be clever**. What do I mean by this? Well: by definition, we know that λ is an eigenvalue of A_G iff $\det(\lambda I - A)$ is zero – in other words, if $\lambda I - A$ is a matrix that does not have full rank.

So: what choices of λ will insure that the matrix

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

doesn't have full rank – i.e. that some of the rows will be dependent on other rows?

One (perhaps screamingly obvious) answer here is $\lambda = -1$, as this choice of λ means that *all* of the rows in $-1I - A$ are identical, and thus that this matrix has rank 1. This then tells us that -1 is an eigenvalue! Furthermore, because the rank of this matrix is 1, the dimension of E_{-1} is $5 - 1 = 4$, and therefore we know that -1 is an eigenvalue with geometric multiplicity 4 (Recall that the geometric multiplicity of an eigenvalue is, by definition, the dimension of its eigenspace.) This then tells us (in turn) that it must have algebraic multiplicity at least 4, because the algebraic multiplicity is always $\geq$ the geometric multiplicity. (Recall as well that the algebraic multiplicity is the number of times an eigenvalue shows up as a root of $p_A(\lambda)$.)

Why do we care? Well: if -1 has algebraic multiplicity ≥ 4, then there can be at most one other eigenvalue left for our matrix A_G, as it's a 5×5 matrix!

Now, we just have to find one last eigenvalue: i.e. one last value of λ such that

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

doesn't have full rank – i.e. that there is some way to combine these rows to get to 0. So: when we're in situations like this, it's sometimes good to look for symmetries in the matrix we can exploit. In this case: notice that every row is kind-of the same: they all have one λ and 4 (-1)'s. So: if we just look at the linear combination of the rows given by adding all of the rows together, we get the vector $(\lambda - 4, \ldots \lambda - 4)$; in specific, setting $\lambda = 4$ means that this combination of the rows is 0! In other words, when $\lambda = 4$, this matrix doesn't have full rank! So $\lambda = 4$ is our last eigenvalue: as well, because -1 has algebraic and geometric multiplicity 4, 4 must have algebraic and geometric multiplicity 1.

We've found all of our eigenvalues. Now, let's find the eigenspaces!

For E_4, this is pretty trivial: we just want to find

$$\begin{aligned} E_4 &= \text{nullspace}(A - 4I) \\ &= \begin{pmatrix} -4 & 1 & 1 & 1 & 1 \\ 1 & -4 & 1 & 1 & 1 \\ 1 & 1 & -4 & 1 & 1 \\ 1 & 1 & 1 & -4 & 1 \\ 1 & 1 & 1 & 1 & -4 \end{pmatrix}. \end{aligned}$$

We know (as shown earlier) that 4 has geometric multiplicty 1; so this nullspace has dimension 1, and we just need to find *any* vector in this nullspace. One obvious candidate – especially given our earlier discussion! – is the vector $(1, 1, 1, 1, 1)$, as this corresponds to taking the sum of all of the columns, which we know is $\mathbf{0}$. As this space is spanned by just one vector, we don't have to orthogonalize it! – but we do have to normalize it, which just involves dividing this vector by its length:

$$\frac{1}{||(1,1,1,1,1)||} \cdot (1,1,1,1,1) = \frac{1}{\sqrt{1^2+1^2+1^2+1^2+1^2}} \cdot (1,1,1,1,1) = \frac{1}{\sqrt{5}}(1,1,1,1,1).$$

We now turn our attention to E_{-1}, which is just the nullspace of $\lambda I - A$:

$$\begin{pmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

What does it mean for a vector to be in the nullspace of this matrix? Well, for a generic vector $(x_1, x_2, x_3, x_4, x_5)$, it just means that

$$\begin{aligned} & \begin{pmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \quad & \begin{pmatrix} (-x_1 - x_2 - x_3 - x_4 - x_5) \\ (-x_1 - x_2 - x_3 - x_4 - x_5) \\ (-x_1 - x_2 - x_3 - x_4 - x_5) \\ (-x_1 - x_2 - x_3 - x_4 - x_5) \\ (-x_1 - x_2 - x_3 - x_4 - x_5) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \quad & x_1 + x_2 + x_3 + x_4 + x_5 = 0. \end{aligned}$$

From here, you could do two things. One option would be to just find four linearly independent vectors that satisfy the above condition that $x_1+x_2+x_3+x_4+x_5 = 0$, hit them with Gram-Schmidt to make them orthogonal, and then normalize. This is completely legit, and will definitely work!

The other option (which I prefer here) is *to be clever!* (Or, at the least, remember that you've already found orthogonal bases for this kind of condition before.) What do I mean here? Well: instead of just picking vectors and making them orthogonal later on, why don't we just be a little careful when we're picking our vectors so that they're orthogonal off the bat?

Specifically: let's start with some simple vector that satisfies the condition $x_1+x_2+x_3+x_4+x_5 = 0$ – say, $(1,-1,0,0,0)$, as that's pretty simple. Now, if we pick a second vector, what do we need to do to make it orthogonal to the first? Well: we need its first two coördinates to be the same, so that when we take their dot product we'll get 0. So, without loss of generality, suppose our vector starts off $(1,1,\_,\_,\_)$. What can we do to make its coördinates sum to 0? Well: $(1,1,-2,0,0)$ works; so let's take that!

For our third vector, we have the same kind of arguments: to make it orthogonal to our first vector, we want its first two coördinates to be equal, and to make it orthogonal to the third vector, we want the third coördinate to also be equal to the first two! So, our vector starts off as $(1,1,1,\_,\_)$; to make its coördinates sum to 0, we can just take $(1,1,1,-3,0)$. Finally, by the exact same arguments, we can pick our fourth vector to be $(1,1,1,1,-4)$ – this gives you a collection of four vectors in our nullspace that are all orthogonal, without any Gram-Schmidt at all! This is sometimes nice to be able to do, if only to save time (and because it's cool.)

Now we just have to normalize these vectors, which yields the following set:

$$\frac{1}{\sqrt{2}}(1,-1,0,0,0),\ \frac{1}{\sqrt{6}}\cdot(1,1,-2,0,0),\ \frac{1}{\sqrt{12}}\cdot(1,1,1,-3,0),\ \frac{1}{\sqrt{20}}\cdot(1,1,1,1,0-4).$$

All we have to do now is use these vectors as the columns of a matrix E, and put A_G's corresponding eigenvalues in the appropriate places in a diagonal matrix D, to conclude that

$$A_G = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T.$$

Therefore, we can calculate $(A_G)^{712}$ by just raising the diagonal matrix in the middle to the 712-th power:

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \begin{pmatrix} 4^{712} & 0 & 0 & 0 & 0 \\ 0 & (-1)^{712} & 0 & 0 & 0 \\ 0 & 0 & (-1)^{712} & 0 & 0 \\ 0 & 0 & 0 & (-1)^{712} & 0 \\ 0 & 0 & 0 & 0 & (-1)^{712} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T$$

$$= \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \begin{pmatrix} 4^{712} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T.$$

So: to find the (1,1)-th entry of the above matrix, we just have to multiply the first row of the matrix E by the first column of $D^{712} \cdot E^T$. The matrix D^{712} just multiplies each row of E^T by the appropriate value on its diagonal, so this column is just

$$\begin{pmatrix} \frac{4^{712}}{\sqrt{5}} \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{20}} \end{pmatrix};$$

therefore, the entry in $(1,1)$ is

$$\begin{aligned} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} \frac{4^{712}}{\sqrt{5}} \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{20}} \end{pmatrix} &= \frac{4^{712}}{5} + \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} \\ &= \frac{4^{712}}{5} + \frac{4}{5}. \end{aligned}$$

Thus, there are $\frac{4^{712}}{5} + \frac{4}{5}$-many paths of length 712 from 1 to itself. (Fun fact: if you replace 712 with any even number m, you get the same result: there are $\frac{4^m+4}{5}$-many paths from 1 to itself of length m. In particular, this proves that $4^m + 4$ is always a multiple of 5 whenever m is even, because you always have to have an whole number of paths! cool, right?)

Our first follow-up question is the following:

Question 3 *Suppose we take our graph and start "walking" on it, as follows:*

- *Start at the vertex 1.*
- *After a minute passes, we walk randomly from the vertex we are at to any of the other four vertices – i.e. if we're at vertex 3, we have a $1/4$-chance of walking to vertex 1, a $1/4$ chance of walking to vertex 2, a $1/4$ chance of walking to vertex 4, and a $1/4$ chance of walking to vertex 5.*

How can we interpret this as a probability matrix? In the long run, where are we likely to be?

Solution. So: if we interpret this as a probability matrix, we have the following setup:

- We have five states $\{1, 2, 3, 4, 5\}$, corresponding to being at any of the five vertices in our graph.
- The probability of going from state j to state i after a step is $1/4$, as long as $i \neq j$; if $i = j$, the probability of this happening is 0, as we always walk somewhere else after each step.

So: this gives us a 5×5 matrix P, where p_{ij} =probability of going from state j to state $i = 1/4$, as long as $i \neq j$. If we write this out, we get

$$\begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix} = \frac{1}{4} \cdot A_G,$$

our matrix from last time!

Now, we want to find the long-term behavior of our graph. How do we do this? Well, remember the following theorem:

Theorem 4 *(Perron-Frobenius) If P is a probability matrix such that either $P^m > 0$ for some value of m, or the graph associated to P is strongly connected (recall: a graph is strongly connected iff there is a way to walk from any vertex to any other vertex), then there is a unique stable vector for P. Furthermore, this stable vector $\mathbf{v}$ represents the eventual long-term behavior of P – in other words, if you take any possible "starting state" represented by some probability vector $\mathbf{x}$, we have that*

$$\lim_{n\to\infty} P^m \cdot \mathbf{x} = \mathbf{v};$$

i.e. if we repeatedly apply P to any starting state, we'll eventually approach the stable vector $\mathbf{v}$.

From inspection, it's clearly possible to go from any vertex of our graph to any other vertex in our graph. Therefore, our graph is strongly connected; thus, Perron-Frobenius applies! Consequently, we can determine the long-term behavior of our graph by finding its stable vector. (Recall: a stable vector for a probability matrix P is just an eigenvector with eigenvalue 1, that's also a probability vector. All probability matrices have at least one stable vector.)

So: we're looking for a vector $\mathbf{v}$ such that $P \cdot \mathbf{v} = \mathbf{v}$. One way we could find this vector is just to look at the nullspace of $P - I$; this will definitely work! However, a faster way is to use our observation that $P = \frac{1}{4}A_G$: this tells us that if $P \cdot \mathbf{v} = \mathbf{v}$, we must have

$$\frac{1}{4}A_G \cdot \mathbf{v} = \mathbf{v}; \Rightarrow A_G \cdot \mathbf{v} = 4\mathbf{v}.$$

So, in other words, we just want an eigenvector for the eigenvalue 4 in A_G! Our earlier work showed that all of these eigenvectors were of the form $c \cdot (1, 1, 1, 1, 1)$, for some constant c; in particular, if we pick $c = 1/5$, we will have $(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$, which is P's unique stable vector.

Therefore, in the long run, we have an equal probability of being at any of the five vertices in our graph – which is what we'd expect, given that we're walking at random and can go anywhere!

Question 5 *Take A_G once more. What is a singular value decomposition for A_G? Find A_G's pseudoinverse, $(A_G)^+$. What is the minimum value taken by the function $||(A_G)^+ \cdot \mathbf{x}||/||\mathbf{x}||$ over $\mathbb{R}^5 - \{\mathbf{0}\}$?*

Solution. So: recall that a **singular value decomposition** for a $m \times n$ matrix 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

$$A = UDV^T.$$

In general, finding one of these is rather hard. However, for our matrix A_G, we've already found one! If we recall problem 1, in particular, we expressed A_G as the product

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T,$$

or EDE^T for short, where E is an orthogonal matrix and D is a diagonal matrix. This is a singular value decomposition – just let $U = E, D = D$, and $V = E$!.

With this done, we now want to find A_G's pseudoinverse. To do this, we recall a theorem from earlier, which said that if a matrix A had UDV^T as its singular value decomposition, its pseudoinverse was given by VD^+U^T, where D^+ was the matrix formed by taking D^T and replacing all of its nonzero values with their reciprocals.

So: if we simply apply this, we have that $(A_G)^+ = ED^+E^T =$

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T.$$

Fun fact – if you multiply A_G and $(A_G)^+$, you'll get $EDE^T \cdot ED^+E^T = ED \cdot D^+E^T = EE^T = I$. So A_G's pseudoinverse is actually its real inverse – i.e. $(A_G)^+ = (A_G)^{-1}$! Cool, right?

Finally, to resolve this last question, we just remember the final theorem we have about singular value decompositions – if we want to find the minimum value of $||(A_G)^+ \cdot \mathbf{x}||/||\mathbf{x}||$ over $\mathbb{R}^5 - \{\mathbf{0}\}$, we just need to find the singular value decomposition of $(A_G)^+$, find the smallest (in terms of absolute value) element in the middle-diagonal part, and pick the column $\mathbf{v}_i$ of V that corresponds to that smallest element. Then $||(A_G)^+ \cdot \mathbf{v}_i||$ will be that minimal value!

So: from above, we can see that $A_G = ED^+E^T$, and that the smallest value on the diagonal of D^+ is 1/4, the entry at (1,1), in terms of absolute value. Therefore, the corresponding column of E that we want is its first column – the vector $(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}})$.

Therefore, we can find the minimum of $||(A_G)^+ \cdot \mathbf{x}||/||\mathbf{x}||$ by looking at the magnitude of the vector

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix}^T \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & \frac{-3}{\sqrt{12}} & \frac{1}{\sqrt{20}} \\ \frac{1}{\sqrt{5}} & 0 & 0 & 0 & \frac{-4}{\sqrt{20}} \end{pmatrix} \cdot \begin{pmatrix} 1/4 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \frac{1}{4} \cdot \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix},$$

which has length 1/4. So the minimum length achievable here is 1/4.