

In this document, if $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$ is an $m \times n$ matrix, $\text{ref}(A)$ is a row-equivalent matrix in row-echelon form using Gaussian elimination with partial pivoting as described in class.

Inner product and orthogonality

Problem

What is the largest possible magnitude of $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$?

Solution

By Cauchy-Bunyakovsky-Schwarz inequality, $|\langle \mathbf{u}_1, \mathbf{u}_2 \rangle| \leq \|\mathbf{u}_1\|_2 \|\mathbf{u}_2\|_2$.

Problem

If $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal, find $\|\mathbf{u}_1 + \mathbf{u}_2\|_2^2$.

Solution

By the properties of the 2-norm and orthogonal vectors,

$$\|\mathbf{u}_1 + \mathbf{u}_2\|_2^2 = \langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{u}_1 + \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_1 \rangle + \cancel{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}^0 + \cancel{\langle \mathbf{u}_2, \mathbf{u}_1 \rangle}^0 + \langle \mathbf{u}_2, \mathbf{u}_2 \rangle = \|\mathbf{u}_1\|_2^2 + \|\mathbf{u}_2\|_2^2,$$

so $\|\mathbf{u}_1 + \mathbf{u}_2\|_2^2 = \|\mathbf{u}_1\|_2^2 + \|\mathbf{u}_2\|_2^2$.

Problem

Define an angle between two vectors $\mathbf{u}_1$ and $\mathbf{u}_2$.

Solution

If $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \|\mathbf{u}_1\|_2 \|\mathbf{u}_2\|_2 \cos(\theta)$, then $\cos(\theta) = \frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\|\mathbf{u}_1\|_2 \|\mathbf{u}_2\|_2}$, so find the inverse cosine of the given ratio. Two vectors

are *collinear* if $\theta = 0 = 0^\circ$ or $\theta = \pi = 180^\circ$ and two vectors are *orthogonal* if $\theta = \frac{\pi}{2} = 90^\circ$ or $\theta = \frac{3\pi}{2} = 270^\circ$.

Problem

A collection of n vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ are not mutually orthogonal and not normalized.

Solution

Apply the Gram-Schmidt algorithm. See the text book, but in essence:

1. For i from 1 to n do
 - a. Set $\mathbf{v}_i \leftarrow \mathbf{u}_i$;
 - b. Subtract off the projection of $\mathbf{v}_i$ onto $\hat{\mathbf{v}}_j$ for each of the previous $i - 1$ normalized vectors:
For j from 1 to $i - 1$ do,
$$\mathbf{v}_i \leftarrow \mathbf{v}_i - \langle \hat{\mathbf{v}}_j, \mathbf{v}_i \rangle \hat{\mathbf{v}}_j$$
 - c. Normalize the i^{th} vector assuming that $\mathbf{v}_i \neq \mathbf{0}$:

$$\text{Set } \hat{\mathbf{v}}_i \leftarrow \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_2}.$$

If the vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ are linearly independent, this will produce n orthonormal vectors.

Linear independence

Problem

Given a collection of n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in $\mathbf{F}^m$, determine if they are linearly independent.

Solution

Create the matrix $V = (\mathbf{v}_1 \cdots \mathbf{v}_n)$ and find the row-equivalent matrix $\text{ref}(A)$ in row-echelon form. If the rank of V equals n , the vectors are linearly independent; otherwise, they are linearly dependent.

Corollary

If $n > m$, the vectors must be linearly dependent, for the maximum rank of V is m .

Problem

The *span* of a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ includes all linear combinations of these vectors, so all vectors of the form

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$$

where $\alpha_1, \dots, \alpha_n \in \mathbf{F}$. Is this a subspace of V ?

Solution

Yes. The sum of two linear combinations of vectors must still be a linear combination of these vectors, and multiplying a linear combination of a set of vectors is still a linear combination of these vectors.

Problem

Given a collection of n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in $\mathbf{F}^m$, find the dimension of and a basis for the span.

Solution

Create the matrix $V = (\mathbf{v}_1 \cdots \mathbf{v}_n)$ and find the row-equivalent matrix $\text{ref}(A)$ in row-echelon form. The dimension of the span is the rank of A . A basis for the span are those columns of A that correspond to columns in $\text{ref}(A)$ that have leading non-zero entries.

Problem

When is a set of vectors a basis for their span?

Solution

A set of vectors forms a basis their span if and only if the vectors are linearly independent.

Problem

Given a set of linearly independent vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$, suppose we apply Gram-Schmidt. Is the span of the orthonormal set identical to the span of the original set?

Solution

Yes.

Linear operators

Problem

If $A:U \rightarrow V$, what properties must A have for A to be described as linear?

Solution

$A(\mathbf{u}_1 + \mathbf{u}_2) = A\mathbf{u}_1 + A\mathbf{u}_2$ for all vectors $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $A(\alpha\mathbf{u}) = \alpha A\mathbf{u}$ for all vectors $\mathbf{u} \in U$.

That is, if a vector-space operation is performed first in U after which A is applied is the same as if A is applied and the vector-space operation is then performed in V .

Problem

If $A:U \rightarrow V$ is a linear mapping, which is the domain, the co-domain and the range of A ?

Solution

The domain of A is U and the co-domain of A is V . The range of A is the collection of all images $A\mathbf{u}$ of vectors $\mathbf{u} \in U$.

Problem

If $A:U \rightarrow V$ is a linear mapping, is the range a subspace of V ?

Solution

If $\mathbf{v}_1, \mathbf{v}_2 \in V$ that are in the range of A , there must exist $\mathbf{u}_1, \mathbf{u}_2 \in U$ such that $A\mathbf{u}_1 = \mathbf{v}_1$ and $A\mathbf{u}_2 = \mathbf{v}_2$. By the linear properties of A , $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = \alpha A\mathbf{u}_1 + \beta A\mathbf{u}_2 = A(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2)$, and because U is a vector space, $\alpha\mathbf{u}_1 + \beta\mathbf{u}_2 \in U$, thus, $A(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ is also in the range of A . Thus, the range is a subspace of V .

Problem

If $A:U \rightarrow V$ is a linear mapping, what is the image of $\mathbf{0}_U$?

Solution

As $\mathbf{0}_U = 0\mathbf{u}$ for any vector in U , $A(\mathbf{0}_U) = A(0 \cdot \mathbf{u}) = 0 \cdot A\mathbf{u} = \mathbf{0}_V$, so $A(\mathbf{0}_U) = \mathbf{0}_V$.

Problem

How do we find the matrices associated with each of the row operations?

Solution

Apply the row operation in question to the identity matrix:

Row operation	Physical interpretation	Description	Representation	Effect	Inverse
Adding a multiple of one row onto another	shear	Adding α times Row i onto Row j	$R_{\alpha;i \rightarrow j}$	$r_{j,i} = \alpha$	$R_{-\alpha;i \rightarrow j}$
Swap two rows	reflection	Swapping Rows i and j	$R_{i \leftrightarrow j}$	$r_{i,i} = r_{j,j} = 0$ $r_{i,j} = r_{j,i} = 1$	$R_{i \leftrightarrow j}$
Multiplying a row by a non-zero scalar	scaling	Multiplying Row i by α	$R_{\alpha;i}$	$r_{i,i} = \alpha$	$R_{\frac{1}{\alpha};i}$

Null space and range of finite-dimensional linear mappings

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$, find the dimension of and a basis for the range.

Solution

Given A , find the row-equivalent matrix $\text{ref}(A)$ in row-echelon form. The dimension of the range is the rank of A . A basis for the range are those columns of A that correspond to columns in $\text{ref}(A)$ that have leading non-zero entries.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$, find the dimension and a basis for the null space.

Solution

Given A , find the row-equivalent matrix $\text{ref}(A)$ in row-echelon form. The number of free variables equals the dimension of the null space, and to find a basis for the null space, solve $(A \parallel \mathbf{0}_m)$, which is equivalent to solving $(\text{ref}(A) \parallel \mathbf{0}_m)$, which can be solved using backward substitution.

Problem

Given $A: U \rightarrow V$, if $A\mathbf{u} = \mathbf{v}$ and $A\mathbf{u}_0 = \mathbf{0}_V$, then $A(\mathbf{u} + \alpha\mathbf{u}_0) = \mathbf{v}$.

Solution

By the properties of linearity, $A(\mathbf{u} + \alpha\mathbf{u}_0) = A\mathbf{u} + A(\alpha\mathbf{u}_0) = A\mathbf{u} + \alpha A\mathbf{u}_0 = \mathbf{v} + \alpha\mathbf{0}_V = \mathbf{v} + \mathbf{0}_V = \mathbf{v}$.

Problem

Given $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$, argue that the dimension of the null space plus the dimension of the range always equals n .

Solution

Given A , find the row-equivalent matrix $\text{ref}(A)$ that is in echelon form. Every column in $\text{ref}(A)$ that has a leading non-zero entry adds one dimension to the range, and every column in $\text{ref}(A)$ that does not have a leading non-zero entry adds another free variable, and thus adds one dimension to the null space.

One-to-one and onto

Problem

When is a linear mapping one-to-one?

Solution

When every vector in the range has a unique pre-image.

Problem

When is a linear mapping onto?

Solution

When every vector in the co-domain has at least one pre-image.

Problem

When is a linear mapping one-to-one and onto?

Solution

When every vector in the co-domain has a unique pre-image.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$, what are tests for either one-to-one or onto?

Solution

If $\text{ref}(A)$ has no free variables, A is one-to-one. If there is one or more free variables, A is not one-to-one, it is many-to-one.

If $\text{rank}(A) = m$, the mapping is onto. If $\text{rank}(A) < m$, the image of U is a subspace of V not equal to V .

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$, are there cases when A is not one-to-one or onto?

Solution

If $n > m$, A can never be one-to-one, but it may be onto if $\text{rank}(A) = m$.

If $n = m$, A is either one-to-one and onto, or neither. It cannot be one but not the other.

If $n < m$, A can never be onto, but it may be one-to-one if $\text{rank}(A) = n$.

Matrices

Problem

What are the diagonal entries of an $m \times n$ matrix?

Solution

The diagonal entries of a matrix A are all entries $a_{i,i}$ where $i = 1, \dots, \min\{m, n\}$.

Problem

When is a matrix upper triangular? When is it lower triangular? When is it diagonal?

Solution

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

A matrix is lower triangular if all entries to the right of the diagonal are zero.

A matrix is diagonal if all the entries off of the diagonal are zero. Diagonal matrices are the only matrices that are simultaneously both lower and upper triangular.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ is a permutation matrix, describe its properties.

Solution

An $n \times n$ matrix is a permutation matrix if and only if every row has exactly one 1 and each column has exactly one 1 and all other entries are 0.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ is a permutation matrix, what is the result of $A\mathbf{u}$ for $\mathbf{u} \in \mathbf{F}^n$?

Solution

If $a_{i,j} = 1$, this moves the j^{th} entry of $\mathbf{u}$ to the i^{th} entry.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ is a permutation matrix, what is the inverse.

Solution

The inverse of a permutation matrix is its transpose.

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^m$ and you have the PLU decomposition of A with $A = PLU$, how do you solve $A\mathbf{u} = \mathbf{v}$ for a given target vector $\mathbf{v} \in \mathbf{F}^m$.

Solution

$A\mathbf{u} = PLU\mathbf{u}$, so we are solving $PLU\mathbf{u} = \mathbf{v}$. Multiply both sides by the inverse (transpose) of P to get

$$P^T PLU\mathbf{u} = Id_m LU\mathbf{u} = LU\mathbf{u} = P^T \mathbf{v}.$$

Now, $(LU)\mathbf{u} = L(U\mathbf{u})$, so this is equivalent to solving $L(U\mathbf{u}) = P^T \mathbf{v}$. As $\mathbf{u}$ is unknown, so is $U\mathbf{u}$, so let us represent the unknown $U\mathbf{u}$ by $\mathbf{y}$; that is, $\mathbf{y} = U\mathbf{u}$. Thus, we have the system of linear equations represented by

$$L\mathbf{y} = P^T \mathbf{v}.$$

The augmented matrix of this system of linear equations is $\left(L \mid P^T \mathbf{v} \right)$, and as L is lower triangular, we may use forward substitution to solve for $\mathbf{y}$. Now that we have $\mathbf{y}$, we now are solving the system of linear equations represented by $\mathbf{y} = U\mathbf{u}$, so the augmented matrix of this system of linear equations is $\left(U \mid \mathbf{y} \right)$. As U is upper triangular, we may use backward substitution to find $\mathbf{u}$.

The determinant, the trace and the inverse

Problem

Given $A, B: \mathbf{F}^n \rightarrow \mathbf{F}^n$, argue that $\det(BA) = \det(B) \det(A)$.

Solution

Given a region $R \subseteq \mathbf{F}^n$ with a finite and non-zero volume $\text{vol}(R)$, $A(R)$ is the region comprised of the image of each vector in R , and by definition, $\text{vol}(A(R)) = \det(A) \text{vol}(R)$. Next, if $B(A(R))$ is the region comprised of the images of each vector in $A(R)$, $\text{vol}(B(A(R))) = \det(B) \text{vol}(A(R)) = \det(B) \det(A) \text{vol}(R)$. But $B(A(R)) = (BA)(R)$, and therefore $\det(BA) = \det(B) \det(A)$.

Problem

Given $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ where A is either upper triangular, lower triangular or diagonal, find the determinant of A .

Solution

Multiply the diagonal entries of A .

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$, find the determinant of A .

Solution

If $n = 2$ or 3 , we may use the short-cuts we learned in class. Otherwise, for $n > 3$, given A , find the PLU-decomposition of A . Record the number n_p of row swaps that were required to produce P and multiply the determinant of U by $(-1)^{n_p}$.

Problem

Given $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$, find the trace of A denoted $\text{tr}(A)$.

Solution

The trace of A is the sum of the diagonal entries of A .

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$, approximate $\det(Id + \varepsilon A)$.

Solution

For sufficiently small ε , $\det(Id + \varepsilon A) \approx 1 + \varepsilon \text{tr}(A)$.

Problem

Find and approximate the determinants of $A = \begin{pmatrix} 1.2 & 0.1 \\ -0.2 & 0.9 \end{pmatrix}$ and $B = \begin{pmatrix} 1.2 & 0.1 & 0.3 \\ -0.2 & 0.9 & 0.2 \\ 0.2 & 0.1 & 1.1 \end{pmatrix}$.

Solution

$\det(A) = 1.2 \cdot 0.9 - 0.1 \cdot (-0.2) = 1.1$; and $\begin{pmatrix} 1.2 & 0.1 \\ -0.2 & 0.9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 0.1 \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}$, so $\det(A) \approx 1 + 0.1 \cdot 1 = 1.1$.

$\det(B) = 1.13$; and $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 0.1 \begin{pmatrix} 2 & 1 & 3 \\ -2 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$, so $\det(B) \approx 1 + 0.1 \cdot 2 = 1.2$.

Problem

Show that $Id^{-1} = Id$.

Solution

$IdId = Id$, so $Id^{-1} = Id$.

Problem

Show that if A is invertible, then $A^{-1}A = AA^{-1} = Id$.

Solution

By the properties of operator composition and inverses, $A^{-1} = (A^{-1}A)A^{-1} = A^{-1}(AA^{-1})$, and therefore $AA^{-1} = Id$.

Problem

Show that if A and B are invertible, then $(BA)^{-1} = A^{-1}B^{-1}$.

Solution

Using the properties of the inverse and matrix composition, $(A^{-1}B^{-1})(BA) = A^{-1}(B^{-1}B)A = A^{-1}IdA = A^{-1}A = Id$, and therefore $(BA)^{-1} = (A^{-1}B^{-1})$.

Problem

Show that if A are invertible, then $(A^{-1})^{-1} = A$.

Solution

$(A^{-1})^{-1}A^{-1} = (AA^{-1})^{-1} = Id^{-1} = Id$, and therefore $A = (A^{-1})^{-1}$.

Adjoint of linear mappings

Recall that the adjoint of a linear mapping $A:U \rightarrow V$ is that mapping $A^*:V \rightarrow U$ such that $\langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A^*\mathbf{v} \rangle$ for all $\mathbf{u} \in U$ and all $\mathbf{v} \in V$. For $A:\mathbf{R}^n \rightarrow \mathbf{R}^m$, the adjoint is the transpose, denoted A^T . For $A:\mathbf{C}^n \rightarrow \mathbf{C}^m$, the adjoint is the conjugate transpose.

Problem

If $A:U \rightarrow V$, show that $(A^*)^* = A$.

Solution

By the properties of the adjoint,

$$\langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A^*\mathbf{v} \rangle = \langle (A^*)^*\mathbf{u}, \mathbf{v} \rangle,$$

and therefore $A = (A^*)^*$.

Problem

If $A_1, A_2:U \rightarrow V$, show that $(A_1 + A_2)^* = A_1^* + A_2^*$.

Solution

By the properties of the adjoint and composition of linear mappings,

$$\begin{aligned} \langle \mathbf{u}, (A_1 + A_2)^*\mathbf{v} \rangle &= \langle (A_1 + A_2)\mathbf{u}, \mathbf{v} \rangle = \langle A_1\mathbf{u} + A_2\mathbf{u}, \mathbf{v} \rangle = \langle A_1\mathbf{u}, \mathbf{v} \rangle + \langle A_2\mathbf{u}, \mathbf{v} \rangle \\ &= \langle \mathbf{u}, A_1^*\mathbf{v} \rangle + \langle \mathbf{u}, A_2^*\mathbf{v} \rangle = \langle \mathbf{u}, A_1^*\mathbf{v} + A_2^*\mathbf{v} \rangle = \langle \mathbf{u}, (A_1^* + A_2^*)\mathbf{v} \rangle. \end{aligned}$$

Thus $(A_1 + A_2)^* = A_1^* + A_2^*$.

Problem

If $A:U \rightarrow V$, show that $(\alpha A)^* = \alpha^* A^*$.

Solution

By the properties of the adjoint and composition of linear mappings,

$$\langle \mathbf{u}, (\alpha A)^*\mathbf{v} \rangle = \langle (\alpha A)\mathbf{u}, \mathbf{v} \rangle = \langle \alpha(A\mathbf{u}), \mathbf{v} \rangle = \alpha^* \langle A\mathbf{u}, \mathbf{v} \rangle = \alpha^* \langle \mathbf{u}, A^*\mathbf{v} \rangle = \langle \mathbf{u}, \alpha^*(A^*\mathbf{v}) \rangle = \langle \mathbf{u}, (\alpha^* A^*)\mathbf{v} \rangle.$$

Thus $(\alpha A)^* = \alpha^* A^*$.

Problem

If $A:U \rightarrow V$ and $B:V \rightarrow W$, show that $(BA)^* = AB^*$.

Solution

By the properties of the adjoint and composition of linear mappings,

$$\langle \mathbf{u}, (BA)^* \mathbf{w} \rangle = \langle (BA)\mathbf{u}, \mathbf{w} \rangle = \langle B(A\mathbf{u}), \mathbf{w} \rangle = \langle A\mathbf{u}, B^* \mathbf{w} \rangle = \langle \mathbf{u}, A^* (B^* \mathbf{w}) \rangle = \langle \mathbf{u}, (A^* B^*) \mathbf{w} \rangle.$$

Thus $(BA)^* = A^* B^*$.

Problem

If $A:U \rightarrow U$ and A is invertible, show that $(A^{-1})^* = (A^*)^{-1}$.

Solution

By the properties of the adjoint and composition of linear mappings,

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle Id \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle (AA^{-1})\mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A(A^{-1}\mathbf{u}_1), \mathbf{u}_2 \rangle = \langle A^{-1}\mathbf{u}_1, A^* \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, (A^{-1})^* A^* \mathbf{u}_2 \rangle,$$

but as this is true for all $\mathbf{u}_1$ and $\mathbf{u}_2$, thus $(A^{-1})^* A^* = Id$, so $(A^*)^{-1} = (A^{-1})^*$.

Problem

For $A:U \rightarrow V$, a vector $\mathbf{v}$ is orthogonal to all vectors in the range of A if and only if $\mathbf{v}$ is in the null space of A^* .

Solution

$\mathbf{v}$ is orthogonal to all vectors in the range of $A \leftrightarrow \langle A\mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{u} \in U$

$\leftrightarrow \langle \mathbf{u}, A^* \mathbf{v} \rangle = 0$ for all $\mathbf{u} \in U$

$\leftrightarrow \mathbf{v}$ is in the null space of A^*

Problem

Find the linear combination of vectors $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbf{F}^m$ that best approximates a vector $\mathbf{v} \in \mathbf{F}^m$.

Solution

If $U = (\mathbf{u}_1 \cdots \mathbf{u}_n)$ is such that $\text{rank}(U) = \text{rank}(U \begin{smallmatrix} \vdots \\ \mathbf{v} \end{smallmatrix})$, there is either a unique solution or infinitely many solutions based on whether $\text{ref}(U)$ has zero or more than zero free variables, respectively.

Otherwise, if $\text{rank}(U) < \text{rank}(U \begin{smallmatrix} \vdots \\ \mathbf{v} \end{smallmatrix})$, we must find the least-squares solution by solving $(U^* U \begin{smallmatrix} \vdots \\ U^* \mathbf{v} \end{smallmatrix})$. There is either a unique solution or infinitely many solutions based on whether $\text{ref}(U^* U)$ has zero or more than zero free variables, respectively.

Self adjoint and skew adjoint linear operators

A linear operator $A:U \rightarrow U$ is *self-adjoint* if $A = A^*$. If $A:\mathbf{R}^n \rightarrow \mathbf{R}^n$, we say that A is *symmetric*. If $A:\mathbf{C}^n \rightarrow \mathbf{C}^n$, we say that A is *conjugate symmetric*.

A linear operator $A:U \rightarrow U$ is *skew-adjoint* if $A = -A^*$. If $A:\mathbf{R}^n \rightarrow \mathbf{R}^n$, we say that A is *skew-symmetric*. If $A:\mathbf{C}^n \rightarrow \mathbf{C}^n$, we say that A is *conjugate skew-symmetric*.

Problem

Given a linear operator $A:U \rightarrow U$, show that $A + A^*$ is self-adjoint.

Solution

By the properties of the adjoint and operator addition,

$$\begin{aligned}\langle \mathbf{u}_1, (A + A^*)^* \mathbf{u}_2 \rangle &= \langle (A + A^*) \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A\mathbf{u}_1 + A^*\mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A\mathbf{u}_1, \mathbf{u}_2 \rangle + \langle A^*\mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, A^*\mathbf{u}_2 \rangle + \langle \mathbf{u}_1, (A^*)^* \mathbf{u}_2 \rangle \\ &= \langle \mathbf{u}_1, A^*\mathbf{u}_2 \rangle + \langle \mathbf{u}_1, A\mathbf{u}_2 \rangle = \langle \mathbf{u}_1, (A^* + A) \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, (A + A^*) \mathbf{u}_2 \rangle\end{aligned}$$

and therefore $(A + A^*)^* = A + A^*$, so it is self-adjoint.

Problem

Given a linear operator $A:U \rightarrow U$, show that $A - A^*$ is skew-adjoint.

Solution

By the properties of the adjoint and operator addition,

$$\begin{aligned}\langle \mathbf{u}_1, (A - A^*)^* \mathbf{u}_2 \rangle &= \langle (A - A^*) \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A\mathbf{u}_1 - A^*\mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A\mathbf{u}_1, \mathbf{u}_2 \rangle - \langle A^*\mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, A^*\mathbf{u}_2 \rangle - \langle \mathbf{u}_1, (A^*)^* \mathbf{u}_2 \rangle \\ &= \langle \mathbf{u}_1, A^*\mathbf{u}_2 \rangle - \langle \mathbf{u}_1, A\mathbf{u}_2 \rangle = \langle \mathbf{u}_1, (A^* - A) \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, -(A - A^*) \mathbf{u}_2 \rangle\end{aligned}$$

and therefore $(A - A^*)^* = -(A - A^*)$, so it is skew-adjoint.

Problem

Given a linear operator $A:U \rightarrow U$, show that AA^* is self-adjoint.

Solution

By the properties of the adjoint and operator addition,

$$\begin{aligned}\langle \mathbf{u}_1, (AA^*)^* \mathbf{u}_2 \rangle &= \langle (AA^*) \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle A(A^*\mathbf{u}_1), \mathbf{u}_2 \rangle \\ &= \langle A^*\mathbf{u}_1, A^*\mathbf{u}_2 \rangle + \langle \mathbf{u}_1, (A^*)^* (A^*\mathbf{u}_2) \rangle = \langle \mathbf{u}_1, A(A^*\mathbf{u}_2) \rangle = \langle \mathbf{u}_1, (AA^*) \mathbf{u}_2 \rangle\end{aligned}$$

and therefore $(AA^*)^* = AA^*$, so it is self-adjoint.

Problem

Show that if A is self adjoint, then αA is self adjoint if and only if α is real.

Solution

$\langle \mathbf{u}, (\alpha A)^* \mathbf{v} \rangle = \langle (\alpha A) \mathbf{u}, \mathbf{v} \rangle = \alpha^* \langle A \mathbf{u}, \mathbf{v} \rangle = \alpha^* \langle \mathbf{u}, A^* \mathbf{v} \rangle = \langle \mathbf{u}, \alpha^* (A \mathbf{v}) \rangle = \langle \mathbf{u}, (\alpha^* A) \mathbf{v} \rangle$ and $(\alpha A)^* = \alpha^* A$ if and only if $\alpha = \alpha^*$ which is true if and only if α is real.

Problem

Show that if A is skew adjoint, then αA is skew adjoint if and only if α is real.

Solution

$\langle \mathbf{u}, (\alpha A)^* \mathbf{v} \rangle = \langle (\alpha A) \mathbf{u}, \mathbf{v} \rangle = \alpha^* \langle A \mathbf{u}, \mathbf{v} \rangle = \alpha^* \langle \mathbf{u}, A^* \mathbf{v} \rangle = \langle \mathbf{u}, \alpha^* (A \mathbf{v}) \rangle = \langle \mathbf{u}, (\alpha^* A) \mathbf{v} \rangle$ and $(\alpha A)^* = \alpha^* A$ if and only if $\alpha = \alpha^*$ which is true if and only if α is real.

Problem

If $A: U \rightarrow U$, show that A is the sum of a self-adjoint and a skew-adjoint linear mapping.

Solution

Because $A = \frac{1}{2}(A + A) + \frac{1}{2}(A^* - A^*) = \frac{1}{2}A + \frac{1}{2}A + \frac{1}{2}A^* - \frac{1}{2}A^* = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*)$, and $\frac{1}{2}(A + A^*)$ is self adjoint and $\frac{1}{2}(A - A^*)$ is skew adjoint.

Isometric linear operators

Recall that a linear operator $A:U \rightarrow U$ is isometric if and only if $\|A\mathbf{u}\|_2 = \|\mathbf{u}\|_2$ for all vectors $\mathbf{u} \in U$.

If $A:\mathbf{R}^n \rightarrow \mathbf{R}^n$, then A is a matrix with columns that form an orthonormal set, and the matrix is called *orthogonal*.

If $A:\mathbf{C}^n \rightarrow \mathbf{C}^n$, then A is a matrix with columns that form an orthonormal set, and the matrix is called *unitary*.

Problem

Show that every isometric linear operator is invertible.

Solution

If $A:V \rightarrow V$ is isometric, then $\|A\mathbf{v}\|_2 = \|\mathbf{v}\|_2$ for all vectors, but if A is not invertible, there exists a non-zero $\mathbf{v}$ such that $A\mathbf{v} = \mathbf{0}$. If this was true, $\|A\mathbf{v}\|_2 = \|\mathbf{0}\|_2 = 0 \neq \|\mathbf{v}\|_2$, as $\mathbf{v}$ was assumed to be non-zero. Thus, the matrix is invertible.

Problem

Show that $A:U \rightarrow U$ is isometric if and only if $A^* = A^{-1}$.

Solution

By the properties of isometric linear operators and the adjoint,

$$\begin{aligned} A \text{ is isometric} &\leftrightarrow \|A\mathbf{u}\|_2 = \|\mathbf{u}\|_2 \\ &\leftrightarrow \|A\mathbf{u}\|_2^2 = \|\mathbf{u}\|_2^2 \\ &\leftrightarrow \langle A\mathbf{u}, A\mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle \\ &\leftrightarrow \langle \mathbf{u}, A^* A\mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle \\ &\leftrightarrow A^* A = Id_V \\ &\leftrightarrow A^* = A^{-1} \end{aligned}$$

Problem

If $A:U \rightarrow U$ is isometric, show that αA is isometric if and only if $|\alpha| = 1$.

Solution

Because A is isometric, $\|A\mathbf{u}\|_2 = \|\mathbf{u}\|_2$, but by the properties of the norm, $\|(\alpha A)\mathbf{u}\|_2 = \|\alpha(A\mathbf{u})\|_2 = |\alpha| \|A\mathbf{u}\|_2 = |\alpha| \|\mathbf{u}\|_2$,

and thus $\|(\alpha A)\mathbf{u}\|_2 = |\alpha| \|\mathbf{u}\|_2 = \|\mathbf{u}\|_2$ if and only if $|\alpha| = 1$.

Corollary

If the field associated with U is the reals, A is isometric if and only if $-A$ is isometric.

Problem

If $A:\mathbf{F}^n \rightarrow \mathbf{F}^n$ is isometric, show that the rows of A also form an orthonormal set.

Solution

As $AA^* = A^*A = Id$, the second says that the columns of A form an orthonormal set, but the first says that the conjugates of the rows of A form an orthonormal set, and if the conjugates of the rows of A form an orthonormal set, then so do the rows themselves.

Problem

Argue that a permutation matrix is isometric.

Solution

$A^T A = Id$ if A is interpreted as a real matrix, and if A is interpreted as a complex matrix, as all the entries are real, $A^* A = Id$, so in either case, the inverse is the adjoint, in which case, it is isometric (orthogonal for real matrices and unitary for complex matrices).

Eigenvalues and eigenvectors

Problem

If $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$, find the dimension and a basis for the eigenspace corresponding to a given and known eigenvalue λ .

Solution

Given A and λ , find the row-equivalent matrix $\text{ref}(\lambda Id_n - A)$. The number of free variables equals the dimension of the eigenspace, and to find a basis for the eigenspace, solve $(\lambda Id_n - A \vdots \mathbf{0}_n)$, which is equivalent to solving $(\text{ref}(\lambda Id_n - A) \vdots \mathbf{0}_n)$, which can be solved using backward substitution.

Problem

Given a matrix, find the eigenvalues.

Solution

If a matrix is $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ is upper triangular, lower triangular or diagonal, the diagonal entries of the matrix are the eigenvalues.

If A is none of these, there is an algorithm beyond the scope of this course, the QR algorithm, that will find eigenvalues in a numerically stable manner. For two and three dimensions, $\det(\lambda Id_n - A)$ is always a polynomial in λ with a leading term λ^n , and thus is of degree n . Such a polynomial must have n complex roots, and these roots are the eigenvalues.

Problem

Show that if A^*A has an eigenvalue λ , that eigenvalue is real.

Solution

Suppose $(A^*A)\mathbf{u} = \lambda\mathbf{u}$ for a non-zero vector $\mathbf{u}$. In this case,

$$\lambda \|\mathbf{u}\|_2^2 = \lambda \langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{u}, \lambda \mathbf{u} \rangle = \langle \mathbf{u}, (A^*A)\mathbf{u} \rangle = \langle \mathbf{u}, A^*(A\mathbf{u}) \rangle = \langle A\mathbf{u}, A\mathbf{u} \rangle = \|A\mathbf{u}\|_2^2,$$

and therefore $\lambda = \frac{\|A\mathbf{u}\|_2^2}{\|\mathbf{u}\|_2^2}$ and as both norms are real, so must λ .

Problem

Show that if A is isometric and has an eigenvalue λ , $|\lambda| = 1$.

Solution

Suppose $A\mathbf{u} = \lambda\mathbf{u}$ for a non-zero vector $\mathbf{u}$. By the properties of a norm and isometric linear operators,

$$|\lambda| \|\mathbf{u}\|_2 = \|\lambda\mathbf{u}\|_2 = \|A\mathbf{u}\|_2 = \|\mathbf{u}\|_2,$$

and therefore $|\lambda| = \frac{\|\mathbf{u}\|_2}{\|\mathbf{u}\|_2} = 1$.

Problem

A linear operator $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ is not invertible (that is, it is singular) if and only if $\lambda = 0$ is an eigenvalue of A .

Solution

A is not invertible

if and only if it is not one-to-one,

if and only if the null space is not just $\{\mathbf{0}_n\}$,

if and only if there is a non-zero vector $\mathbf{u}$ such that $A\mathbf{u} = \mathbf{0}_n = 0\mathbf{u}$,

if and only if 0 is an eigenvalue of A .

Problem

Given $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, show that A has no real eigenvectors.

Solution

If $A\mathbf{u} = \lambda\mathbf{u}$, $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 - u_2 \\ u_1 + u_2 \end{pmatrix} = \begin{pmatrix} \lambda u_1 \\ \lambda u_2 \end{pmatrix}$. Therefore, $\begin{pmatrix} (1-\lambda)u_1 - u_2 \\ u_1 + (1-\lambda)u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. The augmented matrix

corresponding to this system of linear equations is $\left(\begin{array}{cc|c} 1-\lambda & -1 & 0 \\ 1 & 1-\lambda & 0 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1-\lambda & 0 \\ 1-\lambda & -1 & 0 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1-\lambda & 0 \\ 0 & -\lambda^2 + 2\lambda - 2 & 0 \end{array} \right)$.

Because $-\lambda^2 + 2\lambda - 2 \leq -1$, it is always true that the only solution to this is $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Thus, the only vector that

is a scalar multiple of itself under multiplication by A is the zero vector, and thus A has no eigenvalues and no eigenvectors.

Problem

Given $A = \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}$, show that A has only one eigenvector corresponding to the eigenvalue $\lambda = 3$.

Solution

As $(3 \cdot Id_2 - A \mid \mathbf{0}_2) = \left(\begin{array}{cc|c} 0 & -2 & 0 \\ 0 & 0 & 0 \end{array} \right)$, there is one free variable v_1 , so the dimension of the eigenspace is 1, and a

basis for this eigenspace is found by solving this, namely, $\alpha_2 = 0$ so $\begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, so $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is the single

eigenvector corresponding to the eigenvalue $\lambda = 3$.

Problem

A linear operator $A: \mathbf{F}^n \rightarrow \mathbf{F}^n$ can under certain circumstances be written as $A = VDV^{-1}$ where V is a matrix composed of the eigenvectors of A . Describe how $A\mathbf{u} = VDV^{-1}\mathbf{u}$ operates?

Solution

If the matrix A has n linearly independent eigenvectors, those eigenvectors form a basis of $\mathbf{F}^n$. Consequently, we may write any vector $\mathbf{u} = V\mathbf{a}$. To find $A\mathbf{u}$, we may either solve this system of linear equations, or we can find the inverse of V to get $V^{-1}\mathbf{u} = \mathbf{a}$. The entry a_k is the coefficient of the k^{th} eigenvector $\mathbf{v}_k$. Multiplying $D\mathbf{a}$ multiplies the k^{th} entry by λ_k , so the k^{th} entry of $DV^{-1}\mathbf{u}$ is $a_k\lambda_k$. Multiplying this vector by the matrix V calculates the linear combination of eigenvectors $\sum_{k=1}^n \lambda_k a_k \mathbf{v}_k$, which is the result of multiplying $A\mathbf{u}$.

Problem

Under which circumstances can we write $A = VDV^{-1}$ as $A = VDV^*$?

Solution

A symmetric matrix $A: \mathbf{R}^n \rightarrow \mathbf{R}^n$ has n real eigenvalues and n orthogonal eigenvectors. We can therefore normalize these eigenvectors and produce an orthogonal matrix V , so $A = VDV^T$. We say that a symmetric matrix is *orthogonally diagonalizable*, as V is an orthogonal matrix.

A normal matrix $A: \mathbf{C}^n \rightarrow \mathbf{C}^n$ has n complex eigenvalues and n orthogonal eigenvectors. We can therefore normalize these eigenvectors and produce a unitary matrix V , so $A = VDV^*$. We say that a normal matrix is *unitarily diagonalizable*, as V is a unitary matrix.

Problem

If we can write $A = VDV^*$, how can we easily calculate $A^{1000}\mathbf{u}$?

Solution

By the properties of isometric linear operators (square matrices) and operator composition (matrix-matrix multiplication), we have

$$\begin{aligned}
 A^{1000}\mathbf{u} &= (VDV^*)^{1000}\mathbf{u} \\
 &= \underbrace{(VDV^*)(VDV^*)\cdots(VDV^*)(VDV^*)}_{1000 \text{ times}}\mathbf{u} \\
 &= V \underbrace{D(V^*V)D(V^*V)\cdots D(V^*V)}_{999 \text{ times}} DV^*\mathbf{u} \\
 &= V \underbrace{DIdIdDV^*\cdots VDId}_{999 \text{ times}} DV^*\mathbf{u} \\
 &= V \underbrace{DDD\cdots D}_{999 \text{ times}} DV^*\mathbf{u} \\
 &= VD^{1000}V^*\mathbf{u}
 \end{aligned}$$

As D is a diagonal matrix with entries λ_k , D^{1000} is a diagonal matrix with entries λ_k^{1000} , an operation that can be performed on a computer with a single function call.

Problem

If $A: \mathbf{R}^n \rightarrow \mathbf{R}^n$ is not symmetric but still diagonalizable, why may we have issues when using a similar technique to calculate $A^{1000}\mathbf{u} = VD^{1000}V^{-1}\mathbf{u}$?

Solution

If we are calculating the inverse, without more information, this operation may be numerically unstable.