§2.1: Matrix Operations

We have several ways to combine functions to make new functions:

- Addition: if f, g have the same domains and codomains, then we can set $(f+g)\mathbf{x} = f(\mathbf{x}) + g(\mathbf{x})$,
- Composition: if the codomain of f is the domain of g, then we can set $(g \circ f)\mathbf{x} = g(f(\mathbf{x}))$,
- Inverse (§2.2): if f is one-to-one and onto, then we can set $f^{-1}(\mathbf{y})$ to be the unique solution to $f(\mathbf{x}) = \mathbf{y}$.

It turns out that the sum, composition and inverse of linear transformations are also linear, and we can ask how the standard matrix of the new function is related to the standard matrices of the old functions.

Notation:

The (i,j)-entry of a matirx A is the entry in row i, column j, and is written a_{ij} or $(A)_{ij}$.

e.g. $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$

The diagonal entries of A are the entries $a_{11}, a_{22}, \ldots$.

A square matrix has the same number of rows as columns. The associated linear transformation has the same domain and codomain.

A diagonal matrix is square matrix whose nondiagonal entries are 0.

e.g. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The identity matrix I_n is the $n \times n$ matrix whose diagonal entries are 1 and whose nondiagonal entries are 0.
It is the standard matrix for the identity transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = \mathbf{x}$.

e.g. $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Addition:

If A, B are the standard matrices for some linear transformations $S, T : \mathbb{R}^n \to \mathbb{R}^m$, then what is $A + B$, the standard matrix of $S + T$?

Proceed column by column:
First column of the standard matrix of $S + T$
$= (S+T)(\mathbf{e_1})$
$= S(\mathbf{e_1}) + T(\mathbf{e_1})$
= first column of A + first column of B.
i.e. $(i,1)$-entry of $A+B$=$a_{i1} + b_{i1}$.

The same is true of all the other columns, so $(A+B)_{ij}$=$a_{ij} + b_{ij}$.

Example: $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \quad A+B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}.$

Scalar multiplication:

If A is the standard matrix for a linear transformation $S : \mathbb{R}^n \to \mathbb{R}^m$, and c is a scalar, then $(cS)\mathbf{x} = c(S\mathbf{x})$ is a linear transformation. What is its standard matrix cA?

Proceed column by column:
First column of the standard matrix of cS
$= (cS)(\mathbf{e_1})$
$= c(S\mathbf{e_1})$
= first column of A multiplied by c.
i.e. $(i,1)$-entry of cA=ca_{i1}.

The same is true of all the other columns, so $(cA)_{ij} = ca_{ij}$.

Example: $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad c = -3, \quad cA = \begin{bmatrix} -12 & 0 & -15 \\ 3 & -9 & -6 \end{bmatrix}.$

Addition and scalar multiplication satisfy some familiar rules of arithmetic:

Let A, B, and C be matrices of the same size, and let r and s be scalars. Then

a. $A + B = B + A$

b. $(A + B) + C = A + (B + C)$

c. $A + 0 = A$

d. $r(A + B) = rA + rB$

e. $(r + s)A = rA + sA$

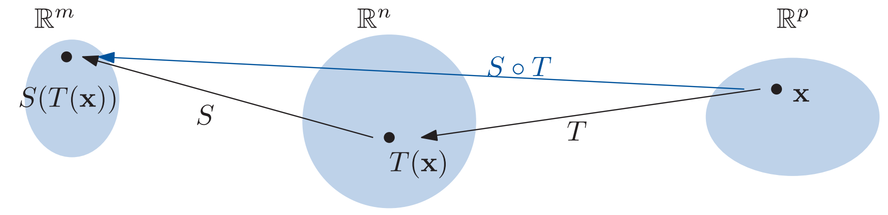
f. $r(sA) = (rs)A$

0 denotes the zero matrix:

$$0 = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

Composition:

If A is the standard matrix for a linear transformation $S : \mathbb{R}^n \to \mathbb{R}^m$ and B is the standard matrix for a linear transformation $T : \mathbb{R}^p \to \mathbb{R}^n$ then the composition $S \circ T$ (T first, then S) is linear.
What is its standard matrix AB?



A is a $m \times n$ matrix,
B is a $n \times p$ matrix,
AB is a $m \times p$ matrix - so the (i, j)-entry of AB cannot simply be $a_{ij}b_{ij}$.

Composition:

Proceed column by column:
First column of the standard matrix of $S \circ T$
$= (S \circ T)(\mathbf{e_1})$
$= S(T\mathbf{e_1})$
$= S(\mathbf{b_1})$
$= A\mathbf{b_1}$, and similarly for the other columns.

So
$$AB = A \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_p \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ A\mathbf{b}_1 & \ldots & A\mathbf{b}_p \\ | & & | \end{bmatrix}.$$

The jth column of AB is a linear combination of the columns of A using weights from the jth column of B.

(i, j)-entry of AB is $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

$$AB = A \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_p \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ A\mathbf{b}_1 & \ldots & A\mathbf{b}_p \\ | & & | \end{bmatrix}.$$

The jth column of AB is a linear combination of the columns of A using weights from the jth column of B.

EXAMPLE: Compute AB where $A = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -3 \\ 6 & -7 \end{bmatrix}$.

Some familiar rules of arithmetic hold for matrix multiplication...

Let A be $m \times n$ and let B and C have sizes for which the indicated sums and products are defined (different sizes for each statement).

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B+C) = AB + AC$ (left - distributive law)
c. $(B+C)A = BA + CA$ (right-distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r
e. $I_m A = A = A I_n$ (identity for matrix multiplication)

... but not all of them:

- Usually, $AB \neq BA$,
- It is possible for $AB = 0$ even if $A \neq 0$ and $B \neq 0$.

Fun application of matrix multiplication:

Consider rotations counterclockwise about the origin.
Rotation through $(\theta + \varphi)$ = (rotation through θ) $\circ$ (rotation through φ).

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

$$= \begin{bmatrix} \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) & -\cos(\theta)\sin(\varphi) - \sin(\theta)\cos(\varphi) \\ \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi) & -\sin(\theta)\sin(\varphi) + \cos(\theta)cos(\varphi) \end{bmatrix}.$$

So $\cos(\theta+\varphi) = \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)$
$\sin(\theta+\varphi) = \cos(\theta)\sin(\varphi) + \sin(\theta)\cos(\varphi)$

Powers:

For a square matrix A, the kth power of A is $A^k = \underbrace{A \dots A}_{k\text{times}}$.

If A is the standard matrix for a linear transformation T, then A^k is the standard matrix for 'applying T k times".

Examples:

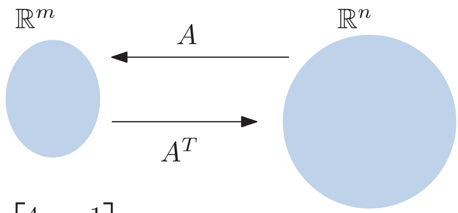
$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}^3 = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 7 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 13 & 14 \\ 21 & 6 \end{bmatrix}.$$

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^3 = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \left(\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \right) = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 27 & 0 \\ 0 & 8 \end{bmatrix}.$$

Exercise: show that $\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}^k = \begin{bmatrix} a_{11}^k & 0 \\ 0 & a_{22}^k \end{bmatrix}$, and similarly for larger diagonal matrices.

Transpose:

The transpose of A is the matrix A^T whose (i,j)-entry is a_{ji}.
i.e. we obtain A^T by "flipping A through the main diagonal".



Example: $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}$, $A^T = \begin{bmatrix} 4 & -1 \\ 0 & 3 \\ 5 & 2 \end{bmatrix}$.

We will be interested in square matrices A such that
$A = A^T$ (symmetric matrix, self-adjoint linear transformation), or
$A = -A^T$ (skew-symmetric matrix), or
$A^{-1} = A^T$ (orthogonal matrix, or isometric linear transformation).

Properties of the transpose:

Let A and B denote matrices whose sizes are appropriate for the following sums and products (different sizes for each statement).

a. $(A^T)^T = A$ (I.e., the transpose of A^T is A)

b. $(A+B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$ (I.e. the transpose of a product of matrices equals the product of their transposes in reverse order.)

Proof: (i,j)-entry of $(AB)^T = (j,i)$-entry of AB

$$\begin{aligned} &= a_{j1}b_{1i} + a_{j2}b_{2i} \cdots + a_{jn}b_{ni} \\ &= (A^T)_{1j}(B^T)_{i1} + (A^T)_{2j}(B^T)_{i2} \cdots + (A^T)_{nj}(B^T)_{in} \\ &= (i,j)\text{-entry of } B^T A^T. \end{aligned}$$

§2.2: The Inverse of a Matrix

Remember from calculus that the inverse of a function $f : D \to C$ is the function $f^{-1} : C \to D$ such that $f^{-1} \circ f =$ identity map on D and $f \circ f^{-1} =$ identity map on C.

Equivalently, $f^{-1}(y)$ is the unique solution to $f(x) = y$.
So f^{-1} exists if and only if f is one-to-one and onto. Then we say f is invertible.

Let T be a linear transformation whose standard matrix is A. From last week:
- T is one-to-one if and only if rref(A) has a pivot in every column.
- T is onto if and only if rref(A) has a pivot in every row.

So if T is invertible, then A must be a square matrix.

Warning: not all square matrices come from invertible linear transformations, e.g. $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Remember from calculus that the inverse of a function $f : D \to C$ is the function $f^{-1} : C \to D$ such that $f^{-1} \circ f =$ identity map on D and $f \circ f^{-1} =$ identity map on C.

Definition: A $n \times n$ matrix A is *invertible* if there is a $n \times n$ matrix C satisfying $CA = AC = I_n$.

Fact: A matrix C with this property is unique:
if $BA = AC = I_n$, then $BAC = BI_n = B$ and $BAC = I_n C = C$ so $B = C$.

The matrix C is called the inverse of A, and is written A^{-1}. So

$$A^{-1}A = AA^{-1} = I_n.$$

A matrix that is not invertible is sometimes called singular.

Remember from calculus that the inverse of a function $f : D \to C$ is the function $f^{-1} : C \to D$ such that $f^{-1} \circ f =$ identity map on D and $f \circ f^{-1} =$ identity map on C.

Equivalently, $f^{-1}(y)$ is the unique solution to $f(x) = y$.

Theorem 5: Solving linear systems with the inverse: If A is an invertible $n \times n$ matrix, then, for each $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof: For any $\mathbf{b}$ in $\mathbb{R}^n$, we have $A(A^{-1}\mathbf{b}) = \mathbf{b}$, so $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution.

And, if $\mathbf{u}$ is any solution, then $\mathbf{u} = A^{-1}(A\mathbf{u}) = A^{-1}\mathbf{b}$, so $A^{-1}\mathbf{b}$ is the unique solution.

In particular, if A is an invertible $n \times n$ matrix, then rref$(A) = I_n$.

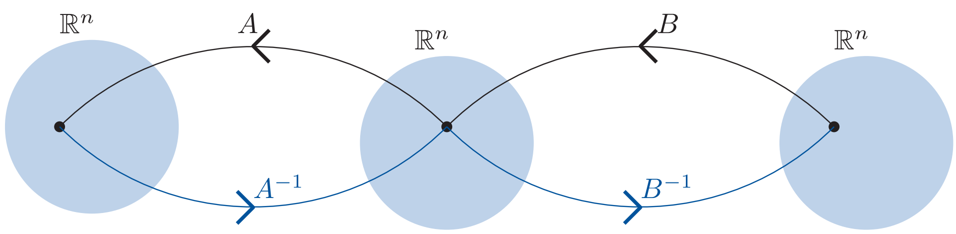
Properties of the inverse:

Suppose A and B are invertible. Then the following results hold:

a. A^{-1} is invertible and $(A^{-1})^{-1} = A$ (i.e. A is the inverse of A^{-1}).

b. AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

c. A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$



Inverse of a 2×2 matrix:

Fact: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

i) if $ad - bc \neq 0$, then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$,

ii) if $ad - bc = 0$, then A is not invertible,

Proof of i):

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\left(\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ba+ab \\ cd-dc & -cb+da \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\left(\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} da-bc & bd-db \\ -ca+ac & -cb+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Proof of ii): next week.

Inverse of a 2×2 matrix:

Example: Let $A = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}$, the standard matrix of rotation about the origin through an angle φ counterclockwise.

$\cos(\varphi)\cos(\varphi) - (-\sin(\varphi))\sin(\varphi) = \cos^2(\varphi) + \sin^2(\varphi) = 1$ so A is invertible, and $A^{-1} = \frac{1}{1}\begin{bmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{bmatrix} = \begin{bmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{bmatrix}$, the standard matrix of rotation about the origin through an angle φ clockwise.

Example: Let $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, the standard matrix of projection to the x_1-axis.

$1 \cdot 0 - 0 \cdot 0 = 0$ so B is not invertible.

Exercise: choose a matrix C that is the standard matrix of a reflection, and check that C is invertible and $C^{-1} = C$.

Inverse of a $n \times n$ matrix:

If A is the standard matrix of an invertible linear transformation T, then A^{-1} is the standard matrix of T^{-1}. So

$$A^{-1} = \begin{bmatrix} | & & | \\ T^{-1}(\mathbf{e}_1) & \dots & T^{-1}(\mathbf{e}_n) \\ | & & | \end{bmatrix}.$$

$T^{-1}(\mathbf{e}_i)$ is the unique solution to the equation $T(\mathbf{x}) = \mathbf{e_i}$, or equivalently $A\mathbf{x} = \mathbf{e_i}$. So if we row-reduce the augmented matrix $[A|\mathbf{e_i}]$, we should get $[I_n|T^{-1}(\mathbf{e}_i)]$. (Remember rref$(A) = I_n$.)

We carry out this row-reduction for all $\mathbf{e_i}$ at the same time:

$$[A|I_n] = \left[\begin{array}{c|ccc} A & | & & | \\ & \mathbf{e}_1 & \dots & \mathbf{e}_n \\ & | & & | \end{array}\right] \xrightarrow{\text{row reduction}} \left[\begin{array}{c|ccc} I_n & | & & | \\ & T^{-1}(\mathbf{e}_1) & \dots & T^{-1}(\mathbf{e}_n) \\ & | & & | \end{array}\right] = [I_n|A^{-1}].$$

If A is an invertible matrix, then

$$[A|I_n] \xrightarrow{\text{row reduction}} [I_n|A^{-1}].$$

EXAMPLE: Find the inverse of $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 1 & 0 & 4 \end{bmatrix}$.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 3 & -1 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1/3 & 0 & 1/3 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4/3 & 0 & -1/3 \\ 0 & 1 & 0 & -5/3 & 1 & -1/3 \\ 0 & 0 & 1 & -1/3 & 0 & 1/3 \end{array}\right]$$

We showed that, if A is invertible, then $[A|I_n]$ row-reduces to $[I_n|A^{-1}]$. The converse is also true:

Fact: If $[A|I_n]$ row-reduces to $[I_n|C]$, then A is invertible and $C = A^{-1}$.

Proof:

If $[A|I_n]$ row-reduces to $[I_n|C]$, then $\mathbf{c}_i$ is the unique solution to $A\mathbf{x} = \mathbf{e}_i$, so $AC\mathbf{e}_i = A\mathbf{c}_i = \mathbf{e}_i$ for all i, so $AC = I_n$.

Also $[C|I_n]$ row-reduces to $[I_n|A]$. So $\mathbf{a}_i$ is the unique solution to $C\mathbf{x} = \mathbf{e}_i$, so $CA\mathbf{e}_i = C\mathbf{a}_i = \mathbf{e}_i$ for all i, so $CA = I_n$.

In particular: an $n \times n$ matrix A is invertible if and only if rref$(A) = I_n$.

Also equivalent: rref(A) has a pivot in every row and column.

For a square matrix, having a pivot in each row is the same as having a pivot in each column.

§2.3: Characterisations of Invertible Matrices

For a square $n \times n$ matrix A, the following are equivalent:

- A is invertible.
- rref(A) $= I_n$.
- rref(A) has a pivot in every row.
- rref(A) has a pivot in every column.

Theorem 8 (The Invertible Matrix Theorem)

Let A be a square $n \times n$ matrix. The the following statements are equivalent (i.e., for a given A, they are either all true or all false).

a. A is an invertible matrix.
b. A is row equivalent to I_n.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I_n$.
k. There is an $n \times n$ matrix D such that $AD = I_n$.
l. A^T is an invertible matrix.

follows from ex. 1a from Monday

ex. 1b from Monday

Important consequences:

- A set of n vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$ if and only if the set is linearly independent (h $\Leftrightarrow$ e).
- If A is a $n \times n$ matrix and $A\mathbf{x} = \mathbf{b}$ has a unique solution for some $\mathbf{b}$, then $A\mathbf{x} = \mathbf{c}$ has a unique solution for all $\mathbf{c}$ in $\mathbb{R}^n$ ($\sim$ d $\Longrightarrow \sim$ g).
- If A is a $n \times n$ matrix and $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution, then there is a $\mathbf{b}$ in $\mathbb{R}^n$ for which $A\mathbf{x} = \mathbf{b}$ has no solution (not d $\Longrightarrow$ not g).
- A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one if and only if it is onto (f $\Leftrightarrow$ i).

Other applications:

Example: Is the matrix $\begin{bmatrix} 1 & -1 & 4 \\ 2 & 0 & 8 \\ 3 & 8 & 12 \end{bmatrix}$ invertible?

Answer: No: the third column is 4 times the first column, so the columns are not linearly independent. So the matrix is not invertible (not e $\Longrightarrow$ not a).

a. A is invertible $\Leftrightarrow$ l. A^T is invertible. (Proof: Check that $(A^T)^{-1} = (A^{-1})^T$.)

This means that the statements in the Invertible Matrix Theorem are equivalent to the corresponding statements with "row" instead of "column", for example:

- The columns of an $n \times n$ matrix are linearly independent if and only if its rows span $\mathbb{R}^n$ (e $\Leftrightarrow$ h^T). (This is in fact also true for rectangular matrices.)
- If A is a square matrix and $A\mathbf{x} = \mathbf{b}$ has a unique solution for some b, then the rows of A are linearly independent ($\sim$d $\Longrightarrow$ e^T).

Advanced application (important for probability):

Let A be a square matrix. If the entries in each column of A sum to 1, then there is a vector $\mathbf{v}$ such that $A\mathbf{v} = \mathbf{v}$.

Hint: $(A - I)^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{0}$.