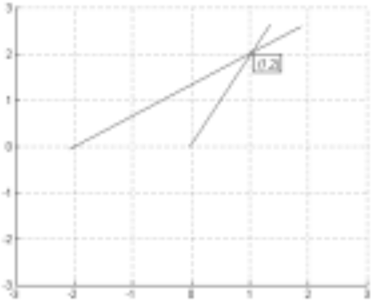
Introduction to Machine Learning - CS725
Instructor: Prof. Ganesh Ramakrishnan

Overview of Linear Algebra

Solving Linear Equation: Geometric View

- Simple example of two equations and two unknowns x and y to be found: $2x - y = 0$ and $-x + 2y = 3$, and in general, $Ax = b$
- One view: Each equation is a straight line in the xy plane, and we seek the point of intersection of the two lines (Fig. 2)



- Challenging in Higher Dimensions!

Three Different Views

- Linear algebra, shows us three different ways of view solutions if they exist):
  1. A direct solution to $Ax = b$, using techniques called elimination and back substitution. [Procedural] [Gauss]
  2. A solution by "inverting" the matrix A, to give the solution $\mathbf{x} = A^{-1}\mathbf{b}$. [Analytical]
  3. A vector space solution, by looking at notions called the column space and nullspace of A. [Conceptual]

Vectors and Matrices

A pair of numbers represented by a *two-dimensional column vector*:

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

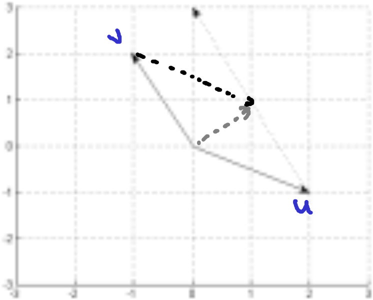
Vector operations: *scalar multiplication* and *vector addition*:
If $\mathbf{v} = (-1, 2)$, then what is $\mathbf{u} + \mathbf{v}$?

Vectors and Matrices

A pair of numbers represented by a *two-dimensional column vector*:

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Vector operations: *scalar multiplication* and *vector addition*:
If $\mathbf{v} = (-1, 2)$, then what is $\mathbf{u} + \mathbf{v}$?



Vectors and Matrices (contd)

- Can be visualised as the diagonal of the parallelogram formed by $\mathbf{u}$ and $\mathbf{v}$
- Any point on the plane containing the vectors $\mathbf{u}$ and $\mathbf{v}$ is some linear combination $a\mathbf{u} + b\mathbf{v}$,
- Space of all linear combinations is simply the full two-dimensional plane ($\Re^2$) containing $\mathbf{u}$ and $\mathbf{v}$
- Similarly, vectors generated by linear combinations of 2 points in a three-dimensional space form some "subspace" of the vector space $\Re^3$
- The space of linear combinations $a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$ could fill the entire three-dimensional space.

Solving Linear Systems: Linear Algebra View

Recap the two equations:

$$2x - y = 0$$

$$-x + 2y = 3$$

And now see their "vector" form:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \tag{1}$$

Solutions as linear combinations of vectors: That is, is there some linear combination of the column vectors $[2, -1]$ and $[-1, 2]$ that gives the column vector $[0, 3]$?

Solving Linear Systems: Linear Algebra View

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

is a 2×2 (Coefficient) Matrix' - a rectangular array of numbers. Further, if

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Then, the matrix equation representing the same linear combination is:

$$A\mathbf{x} = \mathbf{b} \tag{2}$$

A 3×3 Case

$$2x - y = 0$$
$$-x + 2y - z = -1$$
$$-3y + 4z = 4$$

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

Find values of x, y and z such that:

$$x(\text{column 1 of } A) + y(\text{column 2 of } A) + z(\text{column 3 of } A) = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

It is easy to see now that the solution we are after is the solution to the matrix equation $A\mathbf{x} = b$:

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

What about insolvable systems?

It may be the case that for some values of A and b, no values of x, y and z would solve $A\mathbf{x} = b$:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

?

Solution of Linear Equations by (Gauss) Elimination

$$2x - y = 0$$

$$2(-x + 2y = 3)$$

y = 2

back substitution

Progressively *eliminate* variables from equations: First multiply both sides of the second equation by 2 (leaving it unchanged):

$$-2x + 4y = 6$$

Adding LHS of the first equation to the LHS of this new equation, and RHS of the first equation to the RHS of this new equation (does not alter anything):

$$(-2x + 4y) + (2x - y) = 6 + 0 \quad \text{or} \quad \underline{3y = 6}$$

You can see that x has been "eliminated" from the second equation and the set of equations have been said to be transformed into an *upper triangular* form.

$$\begin{aligned} 2x - y &= 0 \\ 3y &= 6 \end{aligned}$$

$$\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

$\Rightarrow y = 6/3 = 2$. And substituting back y into the first equation, $2x - 2 = 0$ or $x = 1$.

Row Elimination: More illustration

$$x + 2y + z = 2$$
$$3x + 8y + z = 12$$
$$4y + z = 2$$

Coefficient matrix:

make it as "1" (pivot)

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}$$

set to "0"

- **The (2,1) step**: First eliminate x from the second equation $\Rightarrow$ multiply the first equation by a multiplier (a_{21}/a_{11}) and subtract it from the second equation.
- a_{11} is called the *pivot*: Goal is to eliminate x coefficient in the second equation.

RHS, after the first elimination step, is:

$$\mathbf{b}_1 = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix}$$

Row Elimination: More illustration

$$A_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}$$

(3,1) (3,2)

- The **(3,1) step for eliminating** a_{31}**:** Nothing to do, so $A_2 = A_1$
- The **(3,2) step for eliminating** a_{32}**:** a_{22} is the next pivot...

$$A_3 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

- A_3 is called an upper triangular matrix → elements below diagonal are 0

- **Sequence of operations on $A\mathbf{x}$ to get $A_3\mathbf{x}$ $\Rightarrow$ multiplying by a sequence of "elimination matrices"**
- Eg: A_1 and $\mathbf{b}_1$ can be obtained by pre-multiplying A and $\mathbf{b}$ respectively by the matrix E_{21}:

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- This also holds for E_{32} and so on. Make sure and **verify** that you understand Matrix multiplication!
- Multiplying matrices A and B is only meaningful if the number of columns of A is the same as the number of rows of B. That is, if A is an $m \times n$ matrix, and B is an $n \times k$ matrix, then AB is an $m \times k$ matrix.

More on Matrix Multiplication

- Matrix multiplication is "associative"; that is, $(AB)C = A(BC)$
- But, unlike ordinary numbers, matrix multiplication is not "commutative". That is $AB \neq BA$
- Associativity of matrix multiplication allows us to build up a sequence of matrix operations representing elimination.

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

- General rule: If we are looking at n equations in m unknowns, and an elimination step involves multiplying equation j by a number q and subtracting it from equation i, then the elimination matrix E_{ij} is simply the $n \times m$ "identity matrix" I, with $a_{ij} = 0$ in I replaced by $-q$.

Elimination as Matrix Multiplication

- For example, with 3 equations in 3 unknowns, and an elimination step that "multiplies equation 2 by 2 and subtracts from equation 3":

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

- The three elimination steps give:

$$E_{32}E_{31}E_{21}(A\mathbf{x}) = E_{32}E_{31}E_{21}\mathbf{b}$$

which, using associativity is:

$$U\mathbf{x} = (E_{32}E_{31}E_{21})\mathbf{b} = \mathbf{c} \tag{3}$$

with U be the obvious upper triangular matrix

Elimination as Matrix Multiplication

$$U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} 2 \\ 6 \\ -10 \end{bmatrix} \qquad (4)$$

- Just as a single elimination step can be expressed as multiplication by an elimination matrix, exchange of a pair of equations can be expressed by multiplication by a *permutation* matrix. Consider..

$$4y + z = 2$$
$$x + 2y + z = 2$$
$$3x + 8y + z = 12$$

The coefficient matrix A can benefit from permutation! Why?

Elimination as Matrix Multiplication

- No solution exists, if, in spite of all exchanges, elimination results in a 0 in any one of the pivot positions
- Else, we will reach a point where the original equation $A\mathbf{x} = \mathbf{b}$ is transformed into $U\mathbf{x} = \mathbf{c}$
- Final step is *back-substitution*, in which variables are progressively assigned values using the right-hand side of this transformed equation
- Eg: $z = -2$, back-substituted to give $y = 1$, which finally yields $x = 2$.

Matrix Inversion for Solving Linear Equations

- Given $A\mathbf{x} = \mathbf{b}$, we find $\mathbf{x} = A^{-1}\mathbf{b}$, where A^{-1} is called the *inverse* of the matrix.
- A^{-1} is such that $AA^{-1} = I$ where I is the identity matrix.
- Since matrix multiplication does not necessarily commute: If for an $m \times n$ matrix A, there exists a matrix A_L^{-1} such that $A_L^{-1}A = I$, $(n \times n)$, then A_L^{-1} is called the left inverse of A.
- Similarly, if there exists a matrix A_R^{-1} such that $AA_R^{-1} = I$ $(m \times m)$, then A_R^{-1} is called the right inverse of A.
- For square matrices, the left and right inverses are the same:

$$A_L^{-1}(AA_R^{-1}) = (AA_L^{-1})A_R^{-1}$$

- For square matrices, we can simply talk about "the inverse" A^{-1}.
- Do all square matrices have an inverse?

Not Every Square Matrix has an Inverse

- Here is a matrix that is not invertible:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \tag{5}$$

- If A^{-1} exists, the solution will be $\mathbf{x} = A^{-1}\mathbf{b}$ and elimination must also produce an upper triangular matrix with non-zero pivots.
- **Thus, the condition works both ways: if elimination produces non-zero pivots then the inverse exists and otherwise, the matrix is not invertible or singular (verify for (5))**

Not Every Square Matrix has an Inverse

- Here is a matrix that is not invertible:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \quad (5)$$

- If A^{-1} exists, the solution will be $\mathbf{x} = A^{-1}\mathbf{b}$ and elimination must also produce an upper triangular matrix with non-zero pivots.
- **Thus, the condition works both ways: if elimination produces non-zero pivots then the inverse exists and otherwise, the matrix is not invertible or singular (verify for (5))**
- $\Leftrightarrow$ **Matrix will be singular iff its rows or columns are linearly dependent (rank $< n$)**

Not Every Square Matrix has an Inverse

- Here is a matrix that is not invertible:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \tag{5}$$

- If A^{-1} exists, the solution will be $\mathbf{x} = A^{-1}\mathbf{b}$ and elimination must also produce an upper triangular matrix with non-zero pivots.
- **Thus, the condition works both ways: if elimination produces non-zero pivots then the inverse exists and otherwise, the matrix is not invertible or singular (verify for (5))**
- **$\Leftrightarrow$ Matrix will be singular iff its rows or columns are linearly dependent (rank $< n$)**
- **$\Leftrightarrow$ Matrix will be singular iff its "determinant" is 0 and is related to the elimination producing non-zero pivots.**

Vector Spaces

- If a set of vectors $\mathcal{V}$ is to qualify as a "vector space" it should be "closed" under the operations of addition and scalar multiplication.
- Thus, given vectors **u** and **v** in a vector space, all scalar multiples of vectors $a\mathbf{u}$ and $b\mathbf{v}$ are in the space, as is their linear combination $a\mathbf{u} + b\mathbf{v}$.
- If a subset (V_S) of any such space is itself a vector space (that is, (V_S) is also closed under linear combination) then (V_S) is called a subspace of (V).
- Eg: Set of vectors $\Re^2$, $\mathcal{M}$ consisting of all 2×2 matrices
- Set $(\Re^2)^+$ (2-D vectors in the positive quadrant is *not* a vector space.

Eg: $a = -5$ $v = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ but $av \notin (\mathbb{R}^2)^+$

Column Space and Solution to Linear System

- *Column space of A, or $C(A)$:* All possible linear combinations of the columns of A, that produce in effect, all possible $\mathbf{b}$'s
- Is there a solution to $A\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{b} \in C(A)$:
- In the example below, is $C(A)$ the entire 4−dimensional space $\Re^4$? If not, how much smaller is $C(A)$ compared to $\Re^4$?

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$$

→ rank = 2
→ $\dim(C(A)) = 2$

- Equivalently, with $A\mathbf{x} = \mathbf{b}$, for which right hand sides $\mathbf{b}$ does a solution $\mathbf{x}$ always exist?
- Definitely does not exist for every right hand side $\mathbf{b}$, (4 equations in 3 unknowns)

- *Which right hand side* **b** *allows the equation to be solved*

$$Ax = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (6)$$

- Eg: If $\mathbf{b} = 0$, the corresponding solution is $\mathbf{x} = \mathbf{0}$. Or whenever $b \in C(A)$ (such as b being a specific column of A).
- Can we get the same space $C(A)$ using less than three columns of A[1]? In this particular example, the third column of A is a linear combination of the first two columns of A. $C(A)$ is therefore a 2−dimensional subspace of $\Re^4$.
- In general, if A is an $m \times n$ matrix, $C(A)$ is a subspace of $\Re^m$.

→ Yes! first 2 cols

[1] In subsequent sections, we will refer to these columns as *pivot* columns.

Null Space

- The null space $N(A)$, is the space of all solutions to the equation $A\mathbf{x} = 0$. → $b = 0$
- $N(A)$ of an $m \times n$ matrix A is a subspace of $\Re^n$.
- Eg: One obvious solution to the system below is 0 (which will always be $\in N(A)$. Any other solution?

$$A\mathbf{x} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (7)$$

Ans: If rank(A) < no of cols, yes!

Finding elements of $N(A)$

- Since columns of A are linearly dependent, a second solution $\mathbf{x}^* \in N(A)$ is as follows (and so are $c\mathbf{x}^*$ for any $c \in \Re$)

$$\mathbf{x}^* = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \tag{8}$$

→ Can be found using Gauss elimination

The null space $N(A)$ is the line passing through the zero vector [0 0 0] and [1 1 $-$1].

- $N(A)$ is always a vector space → $Ax_1 = 0$ & $Ax_2 = 0$, $A(x_1 + x_2) = 0$
- Two equivalent ways of specifying a subspace.
  1. Specify a bunch of vectors whose linear combinations will yield the subspace.
  2. Specify $A\mathbf{x} = \mathbf{0}$ and any vector $\mathbf{x}$ that satisfies the system is an element of the subspace.
- Set of all solutions to the equation $A\mathbf{x} = \mathbf{b}$ - do NOT form a space?

Independence, Basis, and Rank

- **Independence**: *Vectors* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ *are independent if no linear combination gives the zero vector, except the zero combination. That is,* $\forall c_1, c_2, \ldots, c_n \in \Re$, *such that not all of the* c_i*'s are simultaneously* 0, $\sum_i^n c_i \mathbf{x}_i \neq \mathbf{0}$.
- Eg: $\mathbf{x}$ and $2\mathbf{x}$ are dependent
- *The columns* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *of a matrix* A *are independent if the null-space of* A *is the zero vector. The columns of* A *are dependent only if* $A\mathbf{c} = 0$ *for some* $\mathbf{c} \neq \mathbf{0}$.
- **Space spanned by vectors:** *Vectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *span a space means that the space consists of all linear combinations of the vectors. Thus, the space spanned by the columns* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *is* $C(A)$.
- The **rank** of A $(m \times n)$ is the number of its *maximally* independent columns $\leq n$ and those columns form the **basis** of $C(A)$. In the reduced echelon form, all columns will be pivot columns with no free variables.

Not Every Square Matrix has an Inverse

(Recap) Now specify singularity using N(A)

- Here is a matrix that is not invertible:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \tag{9}$$

- If A^{-1} exists, the solution will be $\mathbf{x} = A^{-1}\mathbf{b}$ and elimination must also produce an upper triangular matrix with non-zero pivots.
- **Thus, the condition works both ways: if elimination produces non-zero pivots then the inverse exists and otherwise, the matrix is not invertible or singular (verify for (5))**
- **$\Leftrightarrow$ Matrix will be singular iff its rows or columns are linearly dependent (rank $< n$)**
- **$\Leftrightarrow$ Matrix will be singular iff its "determinant" is 0 and is related to the elimination producing non-zero pivots.**

Singularity and Null Space

- If A^{-1} exists, the only solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = A^{-1}\mathbf{b}$.
- $\Leftrightarrow$ A is singular iff there are solutions other than $\mathbf{x} = \mathbf{0}$ to $A\mathbf{x} = \mathbf{0}$. → A^{-1} does not exist
- $\Leftrightarrow$ A is singular iff it has a non-~~singular~~ trivial null-space $\mathcal{N}(A)$
- Eg: For A in (5), $\mathbf{x} = [3, -1]$ is a solution to $A\mathbf{x} = \mathbf{0}$.

Computing Solution to Linear System (only example)

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \tag{10}$$

elimination[2] changes $C(A)$ while leaving $N(A)$ intact:

$$A_1 = \begin{bmatrix} [1] & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix} \tag{11}$$

$$U = \begin{bmatrix} [1] & 2 & 2 & 2 \\ 0 & 0 & [2] & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12}$$

Row reduced Echelon Form

$U\mathbf{x} = \mathbf{0}$, which has the same solution as $A\mathbf{x} = \mathbf{0}$

$$x_1 + 2x_2 + 2x_3 + 2x_4 = 0$$

$$2x_3 + 4x_4 = 0$$

- Solution can be described by first separating out the two columns containing the pivots, referred to as *pivot columns* and the remaining columns, referred to as *free columns*.
- Variables corresponding to the free columns are called *free variables*, since they can be assigned any value.
- Variables corresponding to the pivot columns are called *pivot variables*
- Following assignment of values to free variables: $x_2 = 1$, $x_4 = 0 \Rightarrow$ by back substitution, we get the following values: $x_1 = -2$ and $x_3 = 0$.

General Procedure

$Ax=b$

$m \times n$

r=m=n	r=m<n	r=n<m
R=I	R=[I F]	$R=[I\ 0]^T$
Unique solution	Infinitely many solutions	0 or 1 solution

from Gauss elimination

under determined

over determined

n-m free vars.

General Procedure

-

Computing the Inverse: From Gauss to Gauss Jordan

- A slight variant, which is invertible:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$$

- How can we determine it's inverse A^{-1}?

$$A^{-1} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{13}$$

The system of equations $AA^{-1} = I$ can be written as:

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- We can solve the two systems to assemble A^{-1}

Gauss Jordan Elimination contd.

$AA^{-1} = I = [b_1\ b_2 \ldots$

- The Guass-Jordan elimination method addresses the problem of solving several linear systems $A\mathbf{x}_i = \mathbf{b}_i$ $(1 \leq i \leq N)$ at once, such that each linear system has the same coefficient matrix A but a different right hand side b_i.
- Key idea: elimination is multiplication by elimination (and permutation) matrices, that transforms a coefficient matrix A into an upper-triangular matrix U:

$$U = E_{32}(E_{31}(E_{21}A)) = (E_{32}E_{31}E_{21})A$$

→ Apply same to $b_1, b_2, b_3 \ldots$

- Now further apply elimination steps until U was transformed into the identity matrix:

$$I = E_{13}(E_{12}(E_{23}(E_{32}(E_{31}(E_{21}A))))) = (E_{13}E_{12}E_{23}E_{32}E_{31}E_{21})A = XA \quad (14)$$

By definition $X = (E_{13}E_{12}E_{23}E_{32}E_{31}E_{21})$ must be A^{-1}.

Illustration of Inversion

- Trick to carry out same elimination steps on two matrices A and B: Create an augmented matrix $[A\ B]$ and carry out the elimination on this augmented matrix.
- Gauss-Jordan: perform elimination steps on the augmented matrix $[A\ I]$ (representing the equation $AX = I$) to give the augmented matrix $[I\ A^{-1}]$ (representing the equation $IX = A^{-1}$).

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right] \overset{Row_2 - 2\times Row_1}{\Longrightarrow} \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right] \overset{Row_1 - 3\times Row_2}{\Longrightarrow} \left[\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

Verify that A^{-1} is

$$A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \tag{15}$$

Dealing with Rectangular Matrices

- What if A is not a square matrix but rather a rectangular matrix of size $m \times n$, such that $m \neq n$. Does there exist a notion of A^{-1}? The answer depends on the rank of A.
  - If A is full row rank and $n > m$, then AA^T is a full rank $m \times m$ matrix $\Leftrightarrow (AA^T)^{-1}$ exists with $A^T(AA^T)^{-1} = I$ and is therefore called the **right inverse of** A. When the right inverse of A is multiplied on its left, we get the projection matrix $A^T(AA^T)^{-1}A$, which projects matrices onto the row space of A.
  - If A is full column rank and $m > n$, then A^TA is a full rank $n \times n$ matrix $\Leftrightarrow (A^TA)^{-1}$ exists with $(A^TA)^{-1}A^T = I$ and is therefore called the **left inverse of** A. When the left inverse of A is multiplied on its right, we get the projection matrix $A(A^TA)^{-1}A^T$, which projects matrices onto the column space of A.
- Singular Value Decomposition: When A is neither full row rank nor full column rank

Full Column Rank and Invertibility

- If A is a full column rank matrix (that is, its columns are independent), $A^T A$ is invertible.
- We will show that the null space of $A^T A$ is $\{0\}$, which implies that the square matrix $A^T A$ is full column (as well as row) rank is invertible. That is, if $A^T A\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$. Note that if $A^T A\mathbf{x} = \mathbf{0}$, then $\mathbf{x}^T A^T A\mathbf{x} = ||A\mathbf{x}|| = 0$ which implies that $A\mathbf{x} = \mathbf{0}$. Since the columns of A are linearly independent, its null space is $\mathbf{0}$ and therefore, $\mathbf{x} = \mathbf{0}$.